Exploring Task-internal and Task-external Readiness:

The Effects of Topic Familiarity and Strategic Planning in Topic-based Task Performance at Different Proficiency Levels

BEI, Xiaoyue

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

探索任務內與任務外預備

主題熟悉程度和策略性構想對不同英語
水準學習者話題性任務表現的效應研究

貝曉越

英語應用語言學
哲學博士論文

香港中文大學
2010 年 7 月

Thesis Committee

Prof. Peter Skehan (Supervisor)

Prof. Jette Hansen Edwards (Co-Supervisor)

Prof. George Braine (Committee Chair)

Prof. Jane Jackson (Committee Member)

Prof. Jean-Marc Dewaele (External Examiner)

Abstract

Research into task-based language teaching (TBLT) has yielded fruitful results with regards to pre-task and during-task preparation activities, with some consensus being reached in a number of areas. Pre-task planning and task repetition usually give rise to fluency and complexity, whereas on-line planning is likely to help with complexity and accuracy. In general, pre-task planning, task repetition, and on-line planning are all task-external manipulations in which extra preparation time is provided so that learners can focus their attention on improving some performance areas. The present study is an attempt to extend the notion of planning from a task-external to a task-internal perspective.

The design of the study is empirical and quantitative in nature. The effects of strategic planning (task-external) and topic familiarity (task-internal) on participants with different proficiency levels are explored and compared. Eighty L2 English participants (forty medicine majors and forty computer majors from a major university in Hong Kong) performed different experimental tasks, in which topic familiarity types, planning conditions and proficiency levels constitute a 2 × 2 × 2 design. Topic familiarity was realized by giving each participant a natural virus topic and a computer virus topic. The topic that matched the participant's academic training was regarded as "familiar", and visa ,versa. There were two types of planning conditions, namely a non-planning (control) group and a ten-minute strategic planning group. Within each

planning condition, the participants were further dichotomized into two different proficiency groups by a proficiency test that was administered prior to the tasks.

Results showed that oooth topic familiarity and strategic planning help learners with more fluent language, but the effect sizes indicate that strategic planning was more powerful in this regard. Topic familiarity pushed learners for slightly more accurate performance with very significant gains in lexical diversity and lexical sophistication. In contrast, planning was associated with significantly higher complexity and lexical density. Proficiency seemed to be concerned with forms rather than meaning expression as higher proficiency participants always scored higher in accuracy and sometimes in complexity, but not so much in fluency or lexis.

Based on these results, task-readiness is suggested as a theoretical extension to the concept of planning to catch both task-external readiness (different types of planning: rehearsal, strategic planning, and on-line planning) and task-internal readiness (content familiarity, schematic familiarity, and task type familiarity). A general framework of task-readiness is proposed as the basis of theorization of task planning and task familiarity.

Abstract in Chinese

論文摘要

任務型語言教學在任務前和任務中各種預備活動的研究上取得了豐碩成果，並一些領域取得了共識。任務前構想和任務重複通常能提高語言的流利度和複雜度，而線上構想則有助於提高語言的複雜度和準確度。總體而言，任務前構想，任務重複和線上構想皆爲“任務外”控制方式，以提供額外的準備時間令學習者將注意力集中於提高語言表現的某些方面。本研究嘗試將“構想”這一概念從“任務外”延伸至“任務內”角度。

本研究本質上屬於實證和定量研究，通過不同分組來探究和比較策略性構想（任務外角度）和主題熟悉程度（任務內角度）對不同英語水準學習者口語任務表現的影響。80 名英語爲二語的實驗參與者（其中 40 名爲醫學系學生, 40 名爲電腦系學生,来自一间香港的大学）參與了各項各項實驗任務，其中主題熟悉程度、策略性構想和英語水準三個變數組成 2 × 2 × 2 實驗設計。主題熟悉程度以給予每位參與者一個自然病毒話題和一個電腦病毒話題口語任務來實現。和實驗參與者專業背景相符合的話題被當作熟悉話題，而不符合者爲不熟悉話題。實驗參與者被分入兩個構想組：無構想組（控制組）和 10 分鐘構想組。這兩組內部又進一步按英語水準前測結果分爲高水準和中水準組。

實驗結果顯示，熟悉話題和策略性構想都提高了實驗參與者的語言流利度，而效果值提示策略性構想比主題熟悉程度在提高流利度方面效果更強。主

題熟悉程度輕微地提高了準確度，同時非常顯著地提高了辭彙多樣性和辭彙複雜度。相較之下，任務前的策略性構想則提高了語言複雜度和辭彙密度。語言水準似乎與語言句法形式而非語言的意義緊密相關，因爲高水準組總是在語言準確度方面，並有時在語言複雜度上，表現更佳。但是語言水準並沒有在語言流利度和辭彙方面展示顯著影響。

基於此等研究結果，本研究提出以"任務預備"做爲"構想"這一概念的理論延伸以囊括"任務內"預備（包括三種不同形式的構想,即任務排演、任務前策略性構想和任務中線上構想）以及"任務外"預備（包括內容熟悉程度、圖式熟悉程度和任務熟悉程度)。本研究提出了"任務預備"的總體理論框架做爲解釋主題熟悉程度和策略性構想效應的理論化基礎。

Acknowledgements

I always take the four-year Ph.D. studies as the gift of heaven so mercifully bestowed upon me. Once warned by many to be a road less taken inasmuch as it requests not only extraordinary intelligence but unfailing perseverance, the years from 2006 to 2010 at the Chinese University of Hong Kong turned out to be a journey of joy and laughter. The person who made all this happen and to whom I am most indebted is my supervisor, Prof. Peter Skehan. He is a model of profound scholarship that I constantly aspire to become, a humorous and knowledgeable professor in whose classes time elapsed too quickly, and a caring as well as friendly gentleman who went hiking and dinning with us students. Never can I forget the support that he offered in every step in my academic pursuit. Nor can the scenes slip from my memory in which he was guiding me in research designing, proposal writing, conducting pilot studies, going through the data, reading, revising my thesis, and so much more. This unusual route from MA to Ph.D. within four years would never have been possible without his guidance and encouragement. His wisdom has always inspired and will continue to inspire me in my future research and life. My deepest gratitude goes to him, Prof. Peter Skehan.

I am very grateful to my committee members, Prof. Jette Hansen Edwards, Prof. George Braine, and Prof. Jane Jackson for their insightful suggestions. Their contributions at various occasions, especially those at my M.Phil progress report and

Ph.D. qualifying exams, have greatly improved the quality of my research. I owe a special thank-you to Prof. Hansen Edwards who took up the responsibility of supervising me in my last term after Prof. Skehan retired. I am deeply grateful to Prof. Jean-Marc Dewaele for his expertise and wonderful feedback as my external examiner.

My thanks should go to Prof. Gerald Nelson and Prof. Gwendolyn Gong for their care and support throughout my studies. My gratitude also goes to Prof. Joanna Radwanska-Williams who has provided me generous help and encouragement in my study program transfers. My deep appreciation goes to Prof. Terri Tomaszek from Davenport University whose love and guidance over so many years have helped me more than I can express.

I would like to thank all my fellow M.Phil and Ph.D. students, especially Wang Zhan (Jan), for their friendship. Two good friends, Mr. Huang Li-long and Ms. Laura Man Ming Yu, have always stood beside me with their love and care to share my happiness and difficulties.

Last but not the least, I thank my parents for their love and understanding of whatever I chose to do. Their caring voice over the phone every week equipped me with so much strength and impetus to overcome any obstacle that I had encountered.

Table of contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 General background of the study

One distinctive feature of second language (L2) speaking is that most learners speak with effort but still fail to reach the native-like proficiency. A tension between the meaning to be expressed and the appropriate forms to use becomes a major challenge to the development of an L2 learner's abilities to use the target language effectively in real life communication. Language teaching then should aim at encouraging learners to perform tasks which not only lend support to their growing L2 competence in making form-meaning connections, but also create environments resembling the real world.

Bygate (2001) pointed out that communicative language ability in general involves the ability to express ideational, interpersonal and discoursal meanings through the use of formal linguistic resources. L2 development in particular further requires helping learners to achieve the capacity to use resources already available to them. This comment can be taken to imply a gap between the 'potential ' and the 'actual', that is, a gap between existing knowledge in one's long term memory and the ability to retrieve it for immediate use in working memory for a communicative task. Such a phenomenon may be attributed to L2 learners' underdeveloped language proficiency, but on top of this, their limited processing capacity (Skehan, 1998) could also have an important role to play.

Therefore, there is a call for exploring pedagogical tasks which go beyond cultivating underlying structural abilities and into increasing learner's readiness for various communicative needs (Samuda, 2001).

Effort in task-based language teaching (TBLT) research attracts attention not only from researchers, but also educational as well as administrative bodies. The Chinese National English Syllabus for secondary schools (2001), for instance, suggests that teachers should try to implement task-based teaching methods. Not very surprisingly, criticism leveled at TBLT never ceases to exist. One of the major concerns that the opponents (e.g., Bruton 2002; Sheen, 1994; Swan 2005) has raised is that practicing teachers were forced to accept a syllabus based on limited research findings and it was therefore premature to put forth such a syllabus for daily use in a wide range of schools.

The logic behind this opposition to TBLT is: if a syllabus is to be (widely) implemented, it requires 'enough' unequivocal research. Whilst it is almost a plain fact that how much research could be deemed not so 'limited' is at best a subjective judgment, we are quite aware that, unlike hard science, it is notoriously difficult, and for some questions even utopistic, to arrive at undisputed conclusions in a social science (though it should be noted that there emerged a trend towards using meta-analyses to pinpoint general patterns out of a larger set of data, as in Norris and Ortega (2002, 2006) and Skehan, (2009)). In fact, the more conventional teaching methodology that Swan and others advocated was not so much established on a solid research foundation, if compared to TBLT. As in any other science, new findings over time in TBLT are the norm whereas the notion of a perfect end point, especially when it becomes the prerequisite for a syllabus, could be regarded as a monkey's wrench thrown into a developing discipline. Certainly, TBLT should respond to criticisms by proceeding in theory construction, empirical research, and padagoical application, and such an idea lies at the heart of the present study.

1.2 Brief rationale of the study

The past 30 years has seen significant advance in task-based language teaching research (TBLT) in at least three areas: *task characteristics* (e.g., whether the task is structured or unstructured; subjective or objective), *task conditions* (e.g., whether the task allows planning or repetition, and whether it is monologic or interactive), and *learners* (e.g., gender, motivation and proficiency) (Skehan, personal communication, 2007). While changes in learner factors require longitudinal research designs and do not seem necessary or appropriate in every case (e.g., gender), task characteristics and task conditions have been shown to affect performance and potential development (see the literature review in the next chapter), and more importantly, they are subject to pedagogical interventions which provide feasible educational means for both teachers and learners.

The present research looks into one of the variables from each of the three above categories in order to gain a more comprehensive view of oral task performance in a single study: *topic familiarity* is a task characteristic; *strategic* *planning* constitutes one of the task conditions; and *proficiency* is meant to be an important individual difference factor in L2 research. Planning of various types (Ellis, 2005), which involves offering learners additional preparation opportunities, has received great attention in TBLT research in recent years. However, some task-inherent characteristics, such as topic familiarity, appear to be under-represented in the literature. Also, as pointed out by Kawauchi (2005), few TBLT studies have seriously considered learner proficiency levels. This is especially true for learners at more advanced levels (Skehan, 2009). It is hoped that an investigation of these three different aspects in TBLT can not only re-visit the more researched variable of planning, but also shed some light on a less touched-upon task-inherent aspect, namely topic familiarity, and see their impact at different proficiency levels.

1.3 Organization of the dissertation

The present dissertation consists of seven chapters that are closely connected with one another. Chapter 1 'Introduction', the present chapter, briefly outlines the general background and rationale to this study and sketches out a roadmap of the whole thesis. This chapter concludes with research questions and hypotheses that guided the whole research project.

Chapter 2 'Literature Review' discusses previous research from three different perspectives: the theoretical background, the empirical background to the study, and methodology issues. The theoretical background provides a summary of theories in three areas: schema, first language (L1) speech production models with different adaptations for L2 speaking, and different definitions and models for communicative competence. The empirical background critically reviews a wide range of studies concerning topic familiarity, strategic planning, and proficiency respectively, which are the three main research variables in the present study. Such an arrangement is meant to link the theories into practice where schema theory is the basis for topic familiarity, speaking models for strategic planning, and models of linguistic competence for proficiency studies. This chapter also extends its coverage into a variety of measurement issues, including the well-noted fluency, complexity and accuracy measures, as well as the less researched measures, i.e. lexis and formality. The theoretical and empirical grounds provide a detailed account of past findings mainly about the three independent variables, whereas the measurement issues prepare the ground for the dependent variables.

Chapter 3 describes the methodology employed in this study. It offers a detailed account of the participants, the C-test as the proficiency measure for grouping purposes, and the two main tasks (one on a familiar topic and the other on an unfamiliar topic). This chapter also provides an overview of the task design and the actual task implementation processes. All independent variables (the performance measures) are tabulated in detail. What comes after the performance measures are the data processing methods with a coding scheme and a speech sample illustrating the way data in this study were analyzed?

Chapter 4 'Pilot studies' does not appear to be a conventionally independent chapter, but the importance of these pilot studies in this study makes it necessary. It will be emphasized that this Ph.D. study is made up of a macro-study and a series of micro-studies that complement the main study. The first two pilot studies were concerned with validating two set of C-tests in a Hong Kong context, based on which a valid and reliable proficiency measure was developed. The third pilot study was a miniature of the main study, which aimed to test the robustness of the topic choices which constitute the pre-condition of the tasks. The third pilot study also trialed the procedure of the tasks. A number of areas that received revision and improvement after the pilot studies will be reported in this chapter.

Chapter 5 presents a comprehensive description of the findings in this research. A MANCOVA result was dealt with in the first place in order to clarify the influence of disciplines and its potential interaction with proficiency levels. Based on this result, a general picture of task performance was gained from MANOVA statistics. Given the fact that the overall MANOVA results allows us to proceed to examine each individual dependent variable, the results in total words, breakdown fluency, repair fluency, accuracy, syntactic complexity, lexis, and formality are each given a detailed portrait. This chapter also continues with two factor analyses to further explore the relationship between all 26 measures, with some interesting findings distinctive from the literature. To address the well-known Robinson vs Skehan debate about theorizing L2 speaking mentioned in Chapter 2, a section on the relationship between accuracy and complexity is provided.

Chapter 6 'Discussion' relates the findings in the last chapter to the literature with the aim of offering explanations and theorization. The first priority was given to a general framework of "task readiness" developed from a combination of findings in the present study and previous research to form the basis for later discussion. This framework extends the notion of planning to task readiness. Planning can be regarded as a kind of task-external readiness, and constructs such as topic familiarity and task type familiarity (Bygate, 2001) could be viewed as task-internal readiness. Based on this general framework, this chapter discusses the results in terms of the three independent variables, namely topic familiarity, strategic planning and proficiency. Whereas the proposal of task readiness creates a theoretical framework for discussion, a further discussion section towards the end of the chapter looks into the results from a wider perspective and synthesizes the results through a series of questions.

Chapter 7 recaps the most important findings in this study with their significant implications. Certain areas of limitations in this study, such as the unequal proficiency levels among the two disciplines, a lack of qualitative analysis, and the imprecise proficiency level in comparison to other studies, are identified. Based on the analyses of the limitations, several directions are suggested for future research.

1.4 Research Questions and Hypotheses

In light of the scarcity of literature in the influence of topic familiarity in L2 speech, and the inconsistency in the previous research in planning and proficiency, the following research questions guide the present study.

1. What are the effects that topic familiarity exerts on L2 oral performance?
2. What are the effects that strategic planning exerts on L2 oral performance?
3. Will proficiency mediate the effects of topic familiarity and/or planning?
4. Will there be an interaction between/among topic familiarity, planning and proficiency?

Drawing on the previous literature on speech production, task-based instruction and schema theory, the following hypotheses are proposed for the research questions:

The Main Effects:

1. It is hypothesized that participants will produce more fluent language under the familiar condition because familiarity with the topic reduces the need to engage in on-line planning, and thus they will have fewer breakdowns and repairs.
2. It is hypothesized that participants will produce more accurate but not more complex language under the familiar condition because familiarity with the topic frees up the attentional resources and makes them available for a focus on form. However, due to the trade-off effects (Skehan, 1996b), complexity will be affected when accuracy

increases.

3. It is hypothesized that familiarity with the topic will bring about higher lexical density, lexical diversity and lexical sophistication since the possession of professional knowledge can render access to the terminology much easier. In contrast, unfamiliarity may result in wandering around the special lexical inventory, and increase the ratio of function words as well as repeated words.

4. It is hypothesized that planning will lead to more fluent language also because of less need for on-line processing and hence fewer breakdowns and repairs.

5. It is hypothesized that planning will lead to more complex sentences but not more accurate language because given planning, participants will be able to develop more complex ideas which may necessitate more complex language, whereas according to Foster and Skehan (1999), without instructor's direction about form, participants are unlikely to concentrate on forms when planning.

6. It is hypothesized that generally there will be a broad effect of proficiency in these speaking tasks since higher-proficient learners have more linguistic resources than the lower-proficient ones.

The Interactions:

<u>*Topic familiarity × planning*</u>

1. Given hypotheses 1 and 4, it is hypothesized that there will be no significant difference in fluency between non-planners in the familiar condition and the planners in the unfamiliar condition in that the two variables influence the same area of performance.
2. Also in the face of hypothesis 1 and 4, it is hypothesized there should be a significant interaction between topic familiarity and planning in fluency. As a result, planners in the familiar condition may outperform other groups, but the effect size should be much smaller than the simple addition of those of topic familiarity and planning.
3. Following hypothesis 2 and 5, it is hypothesized that planners + the familiar condition will be more accurate than the planners + the unfamiliar condition, and more complex than non-planners + familiar, and both more complex and more accurate than the non-planners in the unfamiliar condition. That is, planning and familiarity may compensate for each other regarding accuracy and complexity. This is all because both planning and topic familiarity release part of the demand on working memory, and help the learners to focus on form and increase their willingness to take risks in forming more complex sentence.

Topic familiarity × proficiency

1. Bearing hypothesis 6 in mind, it is hypothesized that specifically, higher proficiency participants in the unfamiliar condition will still outperform the intermediate participants in the familiar condition in terms of fluency and accuracy. However, participants in the familiar conditions will outdo those in the unfamiliar situation in complexity and lexical density, regardless of their proficiency. That is, one variable can be a stronger predictive factor than the other in different areas of performance.

Proficiency × planning

1. It is hypothesized that intermediate planners can outperform high non-planners in terms of fluency and complexity, but the opposite is predicted for accuracy and lexical density. According to Tavakoli and Skehan (2005), planning exerts strong positive effects on task performance at both high and low proficiency levels. At the same time, accuracy and lexical features depend more on their exemplar-based system (see Skehan, 1998 for details of such a system) where proficiency may play a more important role. That is, planning can help compensate for inadequate proficiency in fluency and complexity. However, proficiency is a better predictor with accuracy and lexical density.

Topic familiarity × planning × proficiency

1. It is hypothesized that on the whole this three-way interaction may not be significant because each main effect may be a strong predictor in a certain area of performance. But specifically:

2. The overall order of effects should be: (+familiar + planning + high) > (+familiar +planning – high) OR (+ familiar – planning + high) OR (– familiar + planning + high) > (+ familiar – planning – high) OR (– familiar + planning – high) OR (– familiar – planning + high) > (- familiarity – planning proficiency).

Chapter 2 Literature Review

This chapter critically reviews past research in relevant areas from three major perspectives: theoretical, empirical, and methodological, so that gaps can be indentified to offer directions for the present study. Within the theoretical and empirical perspectives, the discussion will be organized in terms the three major research variables, namely topic familiarity, planning, and proficiency. For the methodological perspective, fluency, accuracy, complexity, lexis, and formality measures will be discussed.

2.1 Theoretical background

The major theoretical foundations for topic familiarity, strategic planning, and proficiency are schema theory, speaking models, and communicative competence, each outlined below.

2.1.1 Schema theory

Bartlett (1932), a British psychologist, was generally accredited as the first scholar to propose the concept of 'schema' to refer to how past experiences are organized in memory and then influence further recall. He however fell short of an explication on the nature of schema (Nassajia, 2002). Schema theory was developed in the 1970s and 1980s (e.g., Minsky, 1975; Schank, 1982) as a general theoretical framework of the structure and function of prior knowledge in the mind. Piaget (1970) identified the cognitive process necessary for comprehension as assimilation – matching sensory information with existing cognitive structures, or the schemata available at any given time. Brown and Yule (1983) described schemata as 'organized background knowledge which leads us to expect or predict aspects in our interpretation of discourse' (p.248). Schank and Abelson (1977) regarded schema as 'stereotypical knowledge' in daily life. Carrell and Eisterhold (1983) defined schema as the structure of prior (background) knowledge. Eysenck and Keane (1990) believed that schema is a group of concepts organized to represent general knowledge. More recently, Carroll (1999) defined a schema as "a structure in semantic memory that specifies the general or expected arrangement of a body of information (p175)".

The notion of 'schema' also appeared under different names, such as 'frame' by Minsky, 'story grammar' by Rumelhart , 'script' as well as the more advanced forms 'MOPs', 'TOPs', and 'TAUs' by Schank (cited in Baddeley, 1997), and 'plans' by Schank (1982). Whichever names or forms schema appears to be, two characteristics of schema can be summarized based on the above descriptions: schema consists of, at a lower level, background knowledge that one has experience in, and thus becomes familiar with. At a higher level, the knowledge is stored in an organized and therefore structured manner, rendering future access easier and faster, like locating information in a book via the table of contents or index.

Schema theory sheds new light on cognitive psychology as it becomes a useful tool for the interpretation of a wide range of cognitive processes, such as comprehending, inferencing and remembering (Nassaji, 2000). What these definitions of schema have in common is that a schema works as a top-down process where new information in working memory will fit into the existing knowledge from long term memory for comprehension to happen. Such a view was challenged by the 'construction-integration' model of comprehension (Kintsch, 1988, 1998).

According to Kintsch (1988, 1998), the theoretical flaw in the traditional top-down view of schema is that readers/listeners have to activate their schema before they can understand incoming information. A serious question is then raised as to how one decides which schema to use before comprehension happens. As Norris and Phillips (1987) pointed out, "... for something to be information, and not just ink marks on a page, it must be understood. However, for the ink mark to be understood one must, by hypothesis in schema theory, have a schema. This leads to a problem wherein schemata are needed to fill in slots in schemata *ad infinitum*" (p.239). To put it simply, in the top-down process of schema, one must have schema ready to understand information, but in order to know which schema to activate, one must have understood the information.

To break this never-ending cycle, Kintsch (1988, 1998) proposed a construction-integration model of comprehension, taking a bottom-up perspective. This process is first data-driven and then schema-driven, where readers/listeners select an appropriate schema on the basis of linguistic cues available. Comprehension takes place more in an interactive than a simple linear fashion, where an existing schema is activated by textual data to allow interpretation and inferences of information. This model, which was well-researched in L1 reading comprehension, emerged as one of the most accepted theoretical framework for comprehension.

Schematic knowledge is often regarded as prior knowledge, background knowledge, or as Bygate (1996) termed, 'world knowledge', all of which have been operationalized in a variety of ways in previous research: 1) cultural knowledge 2) technical knowledge 3) religious knowledge 4) vocabulary knowledge 5) topic familiarity 6) contextual visuals (see Schmidt-Rinehart, 1994, for examples of specific studies of each type).

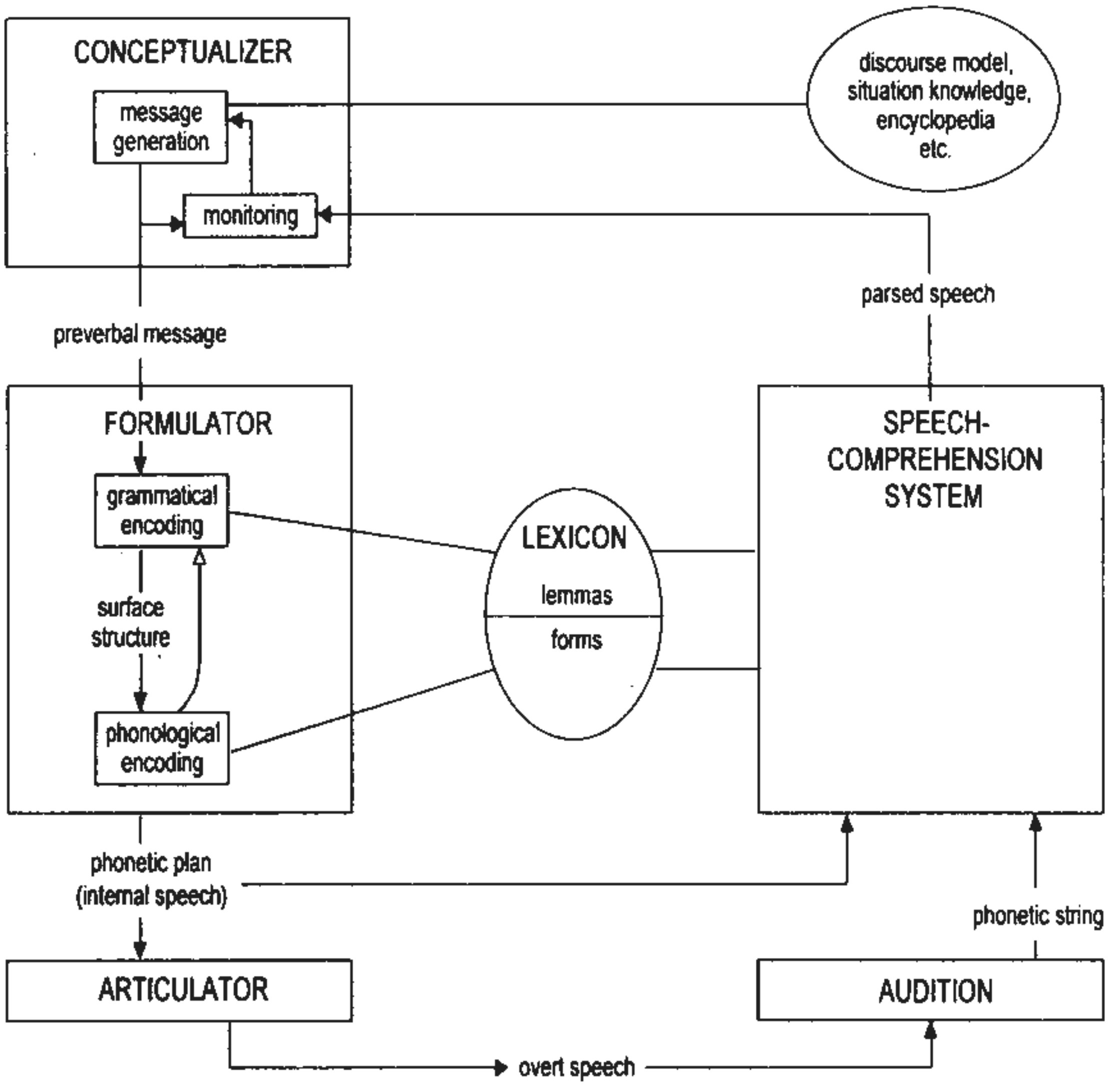
2.1.2 Speaking models

Different models have been proposed to account for L1 speech production (e.g., Garrett, 1981; Levelt, 1989; and Garman, 1990), among which Levelt's (1989) speech processing model is the most influential and the most widely applied theoretical account for task research (e.g., Bygate 2001; Foster and Skehan, 1996; Yuan and Ellis, 2003; Ellis, 2005). Debot (1992), Poulisse (1997), and

Kormos (2006) represented various attempts to adapt Levelt's model for bilingual or L2 speaking.

2.1.2.1 Levelt's models

Figure 1.1 Levelt's (1989) blueprint for the speaker (boxes represent processing components and the circle and the ellipse represents knowledge store)



Levelt's (1989) model is primarily based on findings in the research of speech errors (e.g., tip-of-tongue phenomenon or word substitution) in both normal speakers and speakers with language pathologies (e.g., anomia, a kind of aphasic disturbance in which speakers have difficulties retrieving a word). It is then believed that this model has been established on a robust empirical foundation, with strong explanatory power shown in various studies (e.g., Bygate, 2001; Ellis, 2005; Skehan 2009). Therefore, Levelt's (1989) model is adopted as an important theoretical background to the present study.

Both speech production and comprehension are included in Levelt's (1989) model (see Figure 1), but only the production phase that is relevant to this study will be described here. Three stages, namely conceptualizing the message, formulating the language representations, and articulating the message, were proposed by Levelt (1989). The conceptualization stage is responsible for generating and monitoring the message. It sets the goal of communication and decides on speech acts appropriate for the intended effects. Meanwhile, the conceptualizor also monitors what is to be expressed, what has been expressed, and how to express. At this stage, a general knowledge store, which includes encyclopedic knowledge (about the person's general experience of the world), knowledge about the situation (e.g., the interlocutor/s and the communicative context), as well as information about the discourse record (i.e., what has already been said), is used to generate a message. All this information is then organized into a preverbal message that is not linguistic in nature but contains all information necessary to convert the preverbal message into language. The information perspective of an utterance, its topic, its focus, and the way it would attract the addressee's attention are determined at this point.

This message is then taken over by the formulation stage in which language representation of the preverbal message is established by the retrieval of lexical items from the speaker's mental lexicon. In Levelt's (1989, 1993) model, grammatical and phonological encoding are driven by lexis. Both lexical access procedures and syntactic procedures are applied in grammatical encoding. In the lexicon, each lexical item is specified for semantic and syntactic information (lemmas), and morphological and phonological information (lexemes). When a lemma matches part of the preverbal message, it is retrieved with its syntactic properties to trigger syntactic building procedures. The syntactic properties in lemmas serve to activate the procedural knowledge which works on the syntactic structure of the sentence. Then a string of lemmas is ready for phonological encoding which leads to the selection of specific morphological and phonological forms. A series of phonological segments are activated, and a phonological word is produced. As a result of this stage, a surface structure and a phonetic plan, which is an 'internal speech' in Levelt's term, are constructed, and forwarded to the articulation stage.

This last stage, articulation, 'unfolds and executes (the chunks of the internal speech) as a series of neuromuscular instructions' (Levelt, 1989, p27). The respiratory, the laryngeal, and the super-laryngeal systems are mobilized to carry out the phonetic plan as the result of which overt speech is produced. The articulator is more concerned with a series of physiological processes, which are beyond the scope of this study and will not be discussed in detail here.

Levelt and his colleagues continue to develop and revise the 1989 model (see Levelt, 1993, 1999a, 1999b; Bock and Levelt, 1994; Levelt, Roelofs, and Meyer, 2000). 'Monitoring', for example, was situated inside the conceptualizor in Levelt (1989) to account for internal tracking of errors. From Levelt (1993) onwards, however, the module of monitoring (or called 'self-perception' in Levelt, 1999b) has been given a more independent status outside the conceptualizor to capture the fact that monitoring takes place in other stages (formulator and articulator) as well. Bock and Levelt (1994) followed a 'message – grammatical encoding – phonological encoding – output' language production process, in which grammatical encoding consists of a functional processing component (which involves lexical selection and function assignment) and a positional processing component (which subsumes constituent assembly and inflection). The borderline between the original three stages began to be blurred. More changes were made in Levelt (1999b) in which two principal components were distinguished: the *rhetorical/ semantic/ syntactic system* and the *phonological/phonetic system*. The two systems rely on three knowledge stores to do their processing: the knowledge of external and internal world, the mental lexicon, and the syllabary. The mental lexicon is drawn on by both systems (*lemmas* for grammatical encoding in the first system and *morpho-phonological codes* for morpho-phonological encoding in the second). To sum up, Levelt's models are modular in nature, with different stages being incremental (which allows parallel processing) and lexically-driven.

2.1.2.2 De Bot's (1992) adaptation to Levelt's (1989) model

In an adaptation of Levelt's model for bilingual production, De Bot (1992) suggested that there are two language-specific processing components in the formulation stage, although the two systems are possibly connected in at least some areas. He also believed that the language of the utterance was selected in the Conceptualizer. However, as pointed out by Payne (2002), such a proposal made it difficult to explain fluent code-switching. He explained that "if the preverbal message contains instructions specifying language for the Formulator, then it is not apparent how the speaker would be able to construct parallel speech plans. That is, not only would two Formulators be needed, but also two preverbal messages" (Payne, 2002, p.7). De Bot soon abandoned this proposal to embrace an additional component called 'the Verbalizer' located between the Conceptualizer and the Formulator (De Bot and Schreuder, 1993). The Verbalizer serves as a chunking buffer for the semantic data from the conceptualizer. De Bot and Schreuder (1993) argued that such an addition is necessary in order that the lack of one-to-one correspondence between words and semantic concepts could be explained.

2.1.2.3 Poulisse's (1997) model of bilingual speaking

Poulisse (1997) offered a more comprehensive account of bilingual speaking, with three important differences between L1 and L2 production. First, usually an L2 is not a full-fledged linguistic system. Some lexical items are not yet fully specified for their syntactic, morphological, phonological, and semantic information, which results in slower speech performance in retrieving the right words and also a higher error rate. Second, the L2 production system lacks automaticity in certain aspects of processing. Whereas Levelt (1989) believed that conceptualizor, formulator and articulator function quite automatically in the L1 and can carry out parallel processing, Poulisse (1997) suggested that serial processing is required in lexical retrieving and encoding in L2 production. More attentional resources are then needed in L2 production, which may over-load limited processing capacities (Baddeley and Hitch, 1974; Skehan, 1998) to the detriment of performance. Third, the influence of the L1 is evident in L2 production. This phenomenon involves both conscious code-switching or unconscious L2 transfer. The former happens more in lower proficiency learners who may, for example, have difficulty find a word in the L2. The latter refers to the more systematic cross-linguistic influence of the L1 in L2 production and this can happen without much noticing on the part of the L2 speaker.

Ellis (2005) also believed that:

> ... whereas L1 speakers are able to carry out the process involved in

> formulation and articulation (but not conceptualization) without attention, L2 learners (especially those with limited L2 proficiency) are likely to need to activate and execute their linguistic knowledge through controlled processing. Thus they are likely to experience problems during the formulation and articulation stages, as these processes are demanding on working memory (p.13, parentheses in original).

Taken together, though Levelt's (1989) model provides crucial insights into human speaking processes in general, a different kind of explanation is necessary to account for the special features of L2 speech. Poulisse's (1997) model offers a beneficial supplement to Levelt's model and may help explain various aspects of L2 oral task performance. First of all, Poulisse (1997) believed that the conceptualizor works pretty much the same for both L1 and L2 (except that concepts are tagged for L2), but some aspects of the formulator, such as morphological encoding, are language-specific. It appears that one's background knowledge can always be activated whether in L1 or L2 as far as the conceptualization stage is concerned. Planning time seems more able to help with the formulation stage to compensate for the not yet autonomous lexical encoding processes in L2. Therefore, a more familiar topic can probably lead to more propositions produced and higher fluency due to the easier and faster lexical retrieval; while the opportunity to plan will not only help the speaker to retrieve more content for a message from the conceptualizor, but also probably to formulate the internal speech by matching the lexical items with their appropriate syntactic, semantic, morphological, and phonological properties. This may give rise to faster speech, less hesitation and better syntactic performance. Secondly, Poulisse (1997) and Ellis (2005) both suggested that the level of automaticity depends on proficiency. Higher proficiency could then mean less attentional resource consumption and less 'controlled' processes, which may indicate a higher speech rate and better performance where attentional resources are essential, such as syntactic and lexical encoding.

2.1.2.4 Kormos' (2006) 'Integrated Model of Speech Production'

More recently, Kormos (2006) has, on the basis of extensive reviews on various L1 and L2 speaking models, provided an 'Integrated Model of Speech Production'. Kormos (2006) proposed one single long-term memory store to be shared by both the L1 and the L2, which consists of five components: episodic memory, a hierarchically-structured semantic memory component (including a conceptual base, syntactically-specified lemmas, and morpho-phonologically -specified lexemes), the mental lexicon, the syllabary (automatized gestural scores), and declarative knowledge of L2 syntactic and phonological rules. Unlike the parallel processing nature of L1 speaking (e.g., Levelt, 1989, 1999), the Integrated Model views L2 speaking as basically a serial process due to the lack of automaticity in the formulating stage. Familiar and L2 entries of high frequency occupy a central position in the bilingual lexicon. Similar to Levelt's models, Kormos regarded syntactic processing as lexically-driven, which involves four major stages: activation of the syntactic properties of the lemma corresponding to the first conceptual chunk, phrase and clause-structure building, phonological encoding, and monitoring.

2.1.3 Communicative competence

The well-known dichotomy between competence and performance by Chomsky (1965) held that actual performance is a reflection of one's underlying knowledge system. This account was criticized by Hymes (1972) in that beyond the linguistic domains, the appropriateness of language use *per se* has an abstract element, and is organized, rule-governed, and pervasive (cited in Skehan, 1998). In this sense, a proficiency model involving only formal features such as grammatical, lexical, phonological and idiomatic expressions would not provide sufficient description of a speaker, which is especially true in the case of L2. It is not unusual to see L2 speakers with good syntactic knowledge and a large vocabulary encounter embarrassing communicative breakdowns in real life simply due to the grammatically correct but contextually improper use. A more all-encompassing model on communicative language abilities is then needed for assessing proficiency.

2.1.3.1 Halliday's functional account

From a contrastive perspective to Chomsky, Halliday (1970, 1976, and 1994) argued that grammar is not an independent or autonomous sub-system of one's linguistic ability. The relation of grammar to other 'parts' of the linguistic system is not a part-to-whole relation; rather, it is a symbolic one. Grammar is not independent of meaning, but a resource for creating meaning in the form of wordings. Three components of the human grammatical system were proposed: the ideational, the personal and the textual sub-systems.

The *ideational* component is a grammatical resource for encoding experiences of the world around us and inside us. The *interpersonal* component is used to manage speaker-hearer interaction. It is a grammatical resource for enacting social roles in general, and speech roles in particular, to establish, change, and maintain interpersonal relations. These two sub-systems orient towards two 'extra-linguistic' phenomena: the natural world which is construed in the ideational mode, and the social world which is enacted in the interpersonal mode. The third sub-system, the *textual* component, is concerned with how ideational and interpersonal meanings are organized into coherent discourse that can be shared by speakers and listeners in context. The textual component provides the speakers with strategies for guiding the listeners in his/her interpretations of the text (Matthiessen and Halliday, 1997).

Insights from such a functional account of language abilities are multi-layered: first, it is hard to extract a pure and autonomous 'competence' from 'performance' since all language activities ('performance') are context-dependent. Second, meaning is constructed in context, but not created by linguistic elements (alone). Third, Halliday also touched upon the idea of 'strategy' in the *textual* function of language, thought it is not developed to its full length. All this supports Hymes's idea of communicative competence.

2.1.3.2 The Canale and Swain framework

Two (Canale and Swain, 1980) and three (Canale, 1983) more components are added to the 'linguistic competence' idea originated by Chomsky, making communicative competence a four-dimension model:

Communicative competence	• linguistic • sociolinguistic • discourse • strategic

'Linguistic competence' is similar to Chomsky's formulation of linguistic competence. 'Sociolinguistic' competence derives from Hymes's idea of appropriateness in language use, the individual understanding of social relations, and how language use relates to them. The third component is discourse competence introduced by Canale (1983) to refer to the ability to handle language beyond the sentential level into the discoursal level. The significance of this competence lies in one's understanding of the organization of both spoken and written texts, and the inferences which recover the underlying meaning of what has been said and the connection between utterances. Skehan (1998) pointed out that native-speaker norms are more distinctive in linguistic competence than in the other two just mentioned. Then, it is possible for L2 speakers to develop socio-linguistic and discoursal competences comparable to, or even better than, native speakers', whereas it is much less likely so in the case of linguistic competence. The fourth component in the framework, strategic competence, concerns the compensatory ability to cope with various situations when the other components fail to achieve the intended meaning, or "even to abandon the original meaning and resort to a simpler and more easily achieved goal" (Færch and Kasper, 1983, cited in Skehan, 1998, p.158).

The Canale and Swain proposal extends Hymes's idea of communicative competence into an all-encompassing framework, which provides a more convincing characterization of someone's underlying abilities which can then be related more easily to contexts of actual language use. However, this framework is also limited in its practicability to directly relate the underlying abilities to both performance and processing conditions. In addition, one can hardly generalize language performance from context to context in a systematic way based on this framework. Bachman (1990) developed Canale and Swain's formulation into a more complex, but also more applicable model, which will be discussed below.

2.1.3.3 The Bachman model

Bachman's approach to communicative language competence made advances to earlier models 'in that it attempts to characterize the processes by which the various components interact with each other and with the context in which language use occurs' (Bachman, 1990, p.81). Three major components were involved: language competence, strategic competence, and psychophysiological mechanisms (see Figure 1.2).

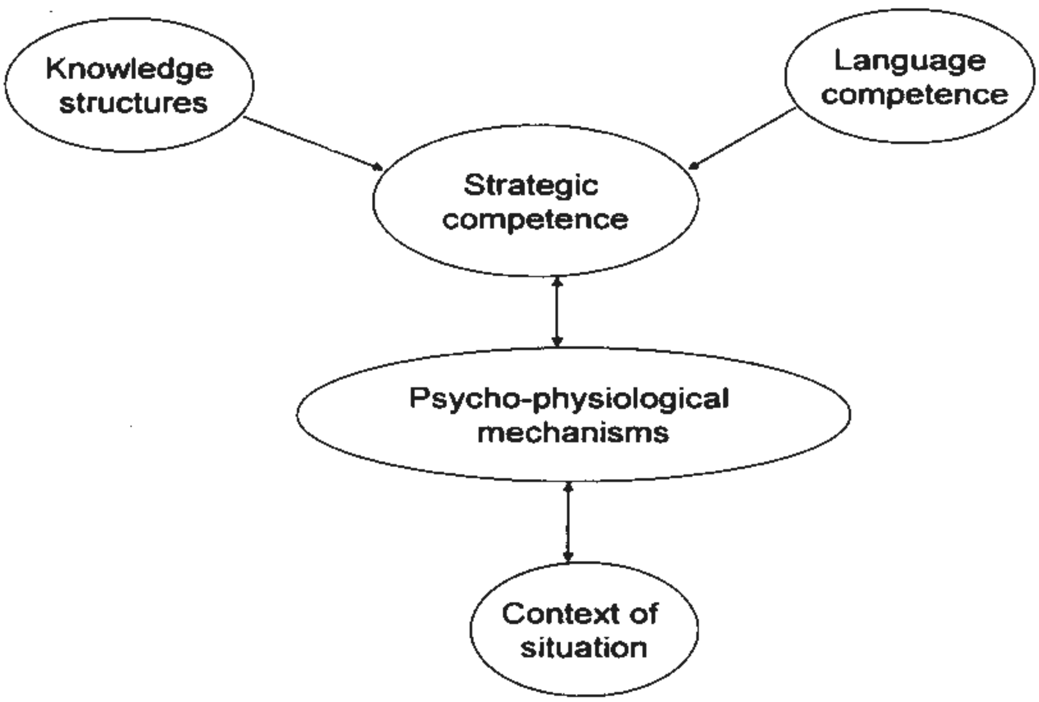
In Bachman's model, the first component, namely language competence, is 'a set of components that are utilized in communication via language' (Bachman, 1990, p.84) As shown in Figure 1.3, language competence consists of two components, organizational competence and pragmatic competence, each of which he further breaks down, with organizational competence covering grammatical and textual competence, and pragmatic competence covering illocutionary and sociolinguistic competence.

The second component is strategic competence which comprises four components: determining communicative goals, assessing communicative resources, planning and executing this communication. It is the mental capacity to implement language competence appropriately in the situation which communication happens, with sociocultural and real world knowledge involved.

The third component, Psychophysiological Mechanisms includes a whole rage of neurological and psychological processes associated with producing and comprehending language.

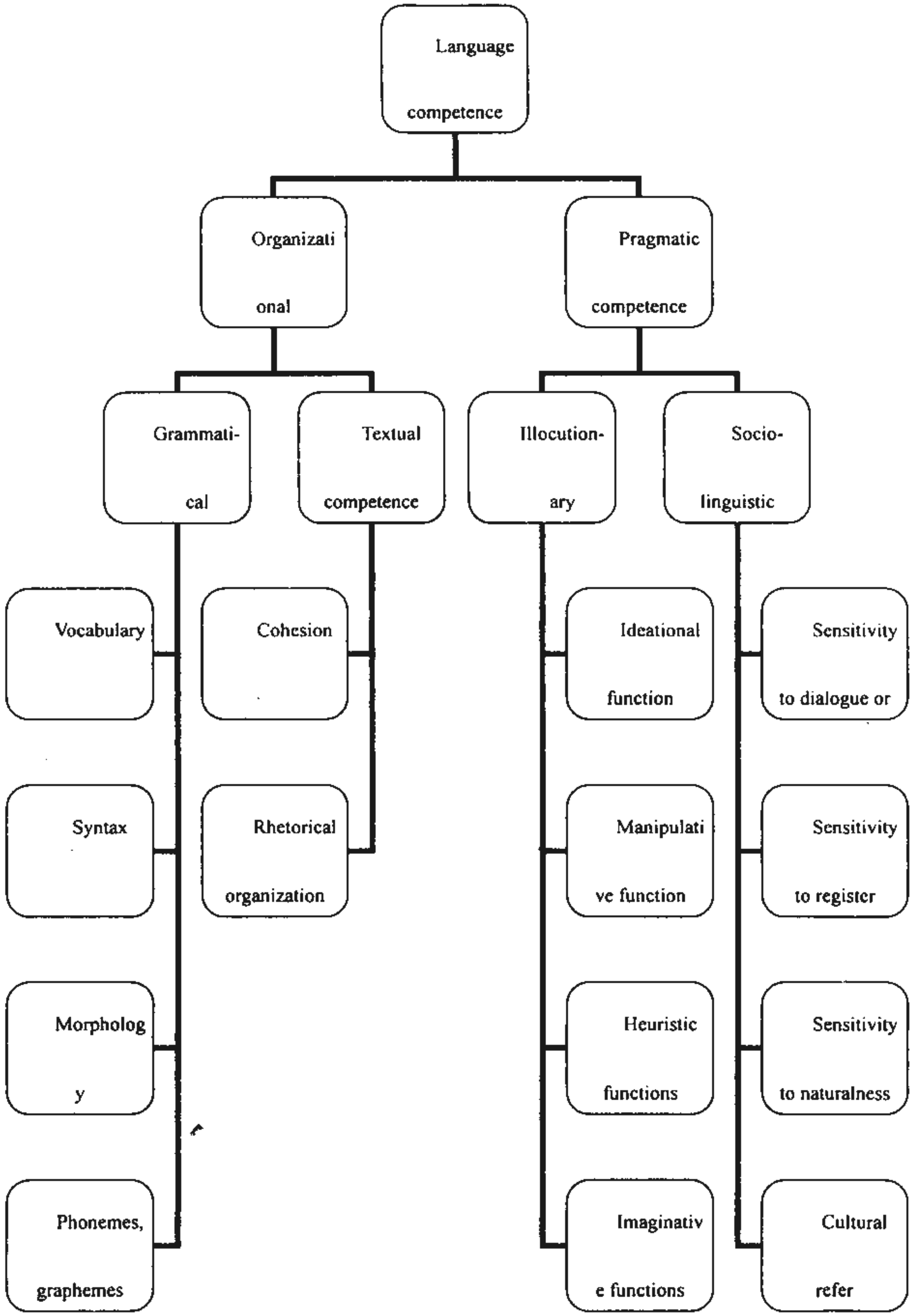
On of the most important contributions Bachman's model made is redefining the relationship between competence and performance, since it now has dynamic

Figure 1.2 Components of communicative language abilities in communicative language use (Bachman, 1990)



qualities (Skehan, 1998). No longer being a compensatory element, strategic competence becomes a central mediating factor between knowledge structures, language competence and the context of situation. This model has several implications related to task performance. First of all, language competence, or proficiency, is an inevitable factor in either measuring task performance or resear-

Figure 1.3 Components of language competence (Bachman, 1990)



-ching the benefits that tasks may have in support of language development. Secondly, the actual enactment of language is influenced not only by language competence, but also one's background knowledge (named as 'knowledge structure' in the model), suggesting a dynamic relationship between one's linguistic knowledge and subject matter knowledge. Thirdly, pre-task planning time may help in at least three out of the four components in strategic competence: goal-setting, assessing and planning, in addition to its functions in retrieving background and linguistic knowledge from long term memory. The three variables in the study find theoretical support from Bachman's (1990) communicative competence model.

2.2 Empirical background

This section is organized in the same order of the three variables as above. However, in the case of proficiency, only studies involving proficiency and also relevant to planning and topic familiarity will be reported to avoid excessive length.

2.2.1 Topic familiarity

Section 2.2.1 consists of three sub-sections, namely empirical research in reading comprehension, listening comprehension, as well as speech production.

2.2.1.1 Reading comprehension

Schema theory provides a general 'slot-filling' framework of reading comprehension. The linguistic cues in a text are held in working memory and get processed to determine which existing schema or schemata to activate. Then the new information is integrated into the schema from long term memory to help make inferences about the meaning of the text. This was called 'default inferencing' in Anderson and Pearson's (1984) term. When reading about a completely new topic where no schema is available, Anderson and Pearson (1984) argued that readers have to rely on logic to make inferences instead of schematic knowledge.

Mixed results were found in research into schematic knowledge in reading comprehension. Schematic knowledge was operationalized in three ways by Carrell (1983): familiarity (reader's own experience with the text content), context (texts with or without a title and a picture) and transparency (plus/minus concrete content words in the texts). She discovered that these three factors all significantly facilitated *native* speakers' reading comprehension, but had little influence on L2 learners. She posited that due to linguistic constraints, L2 learners were not able to exploit background knowledge in their comprehension. Hammadou (1991) took 'sports' and 'AIDS' as two topics for L2 readers themselves to decide the extent to which the topics were familiar or unfamiliar. No positive effect of familiar topic on comprehension was found. However, methodological defects could be found in at least two areas in this study. First, there was no hard control of the 'differentness' between the two topics. Second, complete reliance on students' ratings to decide the extent of familiarity also runs the risk of being misled by participants' subjectivity, especially when the sample size is small. Peretz & Shoham (1990) also reported insignificant results of topic familiarity on reading comprehension.

On the other hand, Shimioda (1993) made use of subject area as + or – familiar in which psychology is + familiar for psychology students, and civil engineering for civil engineering students, and vice versa. This L1 reading experiment reported that in reading comprehension, topic familiarity increases short-term accuracy for recognition questions about concepts. Chang (2006) carried out a reading experiment among learners of Chinese as their L2, employing 'parental love' as the familiar topic due to its universal nature, and folk religions in Taiwan as an unfamiliar topic. The outcome displayed a positive effect of topic familiarity in recall tasks. Topic familiarity also facilitated mental representation for the reading passages. Similarly, Barry and Lazarte (1995), Bügel and Buunk (1996), Chen and Donin (1997), Johnson (1982), and Lee (1986) all found positive effects of topic familiarity on reading comprehension.

One of the two relevant studies published more recently may be Lee (2007) in which the familiar reading text was an expository passage about Korean *Jokpo*, a genealogical record of important historical events and achievements of ancestors (the participants were Korean EFL students) while the unfamiliar passage dealt with the physical process of mummification. The results demonstrated that topic familiarity greatly influenced L2 comprehension, especially in the recall tasks, but had little impact on the acquisition of passive forms (the construction of 'be + V-ed/' as in a regular verb). Coming soon after Lee's (2007) article, Leeser (2007) discovered significant positive effects of topic familiarity on three L2 Spanish tasks, namely comprehension recall, form recognition and tense identification. The novel finding in this study was that topic familiarity also contributed to the learners' ability to make form-meaning connections.

2.2.1.2. Listening comprehension

The processes of schema operation in listening comprehension are quite similar to those in reading as they both belong to the comprehending mode, but the time constraint in listening comprehension imposes additional difficulty on listeners than on readers. The time allowed in listening for the construction process (Kintsch, 1988, 1998) before an appropriate schema can be activated is much shorter, so while L2 readers have the opportunity of going back to the textual data when first-inferencing fails, L2 listeners might encounter troubles at this stage, before any schema is able to take effect. At the same time, schemata might be more important in L2 listening than L2 reading in that unlike readers who might, given less temporal pressure, be able to rely more on linguistic cues bottom-up for meaning construction, listeners probably have no such resource and a schema is crucial for prediction and inferencing in a top-down manner.

In Markham and Latham's (1987) study, a Muslim group, a Christian group and a neutral group that did not have obvious religious beliefs listened to L2 materials about Islamic and Biblical materials to test the effects of +/-familiar background knowledge on listening comprehension. The results are telling: Muslims greatly outperformed Christians in the Islamic passages while at the same time they had fewer distortions of meaning. In the same vein, Christians showed much better understanding of Biblical stories. The control group (called the 'neutral group' in this study) did not display obvious regularity over the different passages. L2 Spanish students in Long (1990) listened to passages in Spanish about Rock Groups and Gold Rushes. In the 'recall protocol' test, these students produced more idea units in the Rock Group than the Gold Rushes condition. Long assumed that current Rock Groups were more familiar to students than historical Gold Rushes, indicating that the more familiar topic helps not only in comprehension, but also retention of the content. Chiang and Dunkel (1992) conducted an experiment with Taiwanese military school students who completed a multiple-choice L2 English listening comprehension task. Confucius was selected as the familiar topic and a foreign place and its people as the unfamiliar one. In this study, the participants scored higher on the familiar topic than the unfamiliar one, but there existed a significant interaction between prior knowledge and test type, which rendered a clear-cut decision on the effect of topic familiarity impossible. Schmidt-Rinehart (1994) determined the variable of topic familiarity according to course syllabus and a questionnaire after the listening task, taking 'Hispanic Universities' as familiar and 'Going for a Walk in the Park' as novel. The results showed that topic familiarity was a powerful factor in listening comprehension. In addition, there was an effect of topic familiarity overriding proficiency whereby students from three course levels all benefited from background knowledge from their textbook. The result was consistent with Douglas (2000) in which he claimed that in tests, background knowledge tends to have a stronger effect on test scores as field specificity increases whereas good language proficiency alone would no longer be sufficient for effective performance.

More recently, Leeser (2004) found that L2 topic familiarity overrode mode in recall tasks. That is, topic familiarity is a stronger predicting factor whether in listening or reading recall. However, in reading comprehension MC tests, there was no significant main effect for topic familiarity or for the covariate, standardized test. A significant interaction was found between mode and familiarity. This revealed the possibility that the effects of topic familiarity might be complicated by mode (reading, listening, and presumably speaking and writing), and question type (recall protocol, MC test and so on).

Taken together, we can conclude that, in general terms, if there is significantly differentiating degree of familiarity between topics, facilitative effects from more familiar topics should be obtained in listening and reading comprehension in comparison to the less familiar ones. As Leeser (2003, cited in

Leeser, 2007) pointed out: the contradictions in the above mentioned reading research could be attributed to a host of methodological issues regarding the assessment of comprehension as well as the operationalization of topic familiarity. There are at least two problems with the above studies. First of all, with few exceptions (e.g., Markham and Latham's, 19[illegible]. Shimioda, 1993), most studies relied on subjective ratings by the research participants. They were obviously lacking in objective criteria to validate the topics. Secondly, none of the studies seriously ensured the comparability of the contrasting topics except for the extent of familiarity to the learners. For example, Hammadou (1991) took 'sports' and 'AIDS' as the task topics. One can argue that, even if a person is a doctor and sports fan at the same time and has the same familiarity with both areas, s/he might still find AIDS more difficult due to its complicated biological nature and a less frequent set of vocabulary associated with it. To sum up, good topics for such a purpose should be comparable as much as possible in every aspect except that learners should have clearly different extent of background knowledge about them.

2.2.1.3. Speech production

Not much research has been conducted on the effects of topic familiarity in L1 speech production; even fewer studies have been conducted in an L2 context. One of the earlier studies, Wiener et al. (1972), claimed that manipulation of the familiarity of the subject matter should have effects on the verbal and nonverbal components of communication. For example, they predicted an increase in gestures which signal uncertainty (e.g., palms up) when the subject discusses less familiar subjects. Good and Butterworth (1980) reported that prior knowledge of a familiar route yielded significantly more fluent language (L1) production than no prior knowledge (describing an unfamiliar route). Li, Williams and Volpe (1995) investigated the effects of topic familiarity in procedural and narrative discourse among L1 aphasic patients and normal subjects. The familiar topics included: 1) going out to dinner, 2) clearing the table after dinner, 3) getting ready for bed, 4) going to the market, and 5) getting a haircut. The unfamiliar topics were 1) going on a mountain, climbing expedition, 2) saddling a horse, 3) making a ceramic vase, 4) making a beanbag chair, and 5) painting a watercolor landscape. They discovered that topic familiarity influenced discourse production in both procedural and story-retell situations. In procedural discourse, a greater number of optional steps were provided with familiar topics. During retelling of familiar topic stories, a greater proportion of action and resolution clauses were included. There were also several studies done more recntely. Bortfeld et al. (2001) found that participants (L1) produced more repeats and restarts, i.e. were less fluent, in describing the less familiar 'tangram' task, but in the more familiar 'children photo' depicting task they produced more fillers. The impact of topic familiarity in L1 oral fluency remains unclear. Merlo and Mansur (2004) had Brazilian Portuguese L1 speakers answer the question 'please tell me what a refrigerator/helicopter is like'. Such familiar (refrigerator) and unfamiliar (helicopter) topics were established by familiarity rating of a list of topics by the participants themselves. They reported that these native speakers produced more propositions about attributes of the familiar object than those of the unfamiliar one whereas there was no significant difference in fluency between the two topics. Banks (2004) investigated a group of about 10-year-old children's L1 narrative performance over a familiar event (a soccer tournament) and an unfamiliar event (soccer tryouts for a more competitive league). It was found that the familiar group did not produce more structurally complex narrative discourse, and a post-test 6 weeks later did not even find improvement in the unfamiliar group though they had completed the task once. Neither was there any significant difference in terms of coherence between the two groups. However, age might be an important factor here since 10-year-old children are not cognitively mature, and therefore differences in complexity and coherence in speech could be hard to spot.

With regard to the influence of topic familiarity in L2 oral production, Chang (1999) conducted an L2 monologic one-way task, finding that topic familiarity resulted in significantly greater fluency (words per error-free T-unit and words per minute), but had no effects on accuracy (error rate per T-unit). However, this study was based on only 6 Taiwanese learners of English and it seems too early to make a strong claim. Another relevant L2 study is Skehan and Foster (1999) in which the structure of a narrative about going to a restaurant can be regarded as familiar to participants while any structure of the narrative about playing golf, unfamiliar.

They found that the structured (schematically familiar) topic could lead to greater fluency because the access to macrostructure of a task reduces the need to engage in lots of mid-task repairs. In addition, the available overall plan allows participants to allocate attention to speech in a more sustained manner. Prior knowledge is part of task complexity along the resource-depleting dimension (Robinson, 2001a), so Robinson would predict that familiarity will lead to increased fluency, but decreased accuracy and complexity. In Robinson's (2001a) study, Japanese learners of English (L2) were instructed to give directions to partners in two map tasks, with one being their university map (familiar) and the other, map of a street area in central Tokyo (unfamiliar). He reported more lexical complexity on the complex/unfamiliar map task whereas there was greater fluency on the simple/familiar task. However, lexical density was employed as the measure for lexical complexity but it was not adjusted for text length in this study. Therefore his 'task complexity leads to lexical complexity' theory was not very persuasively argued.

It appears that, unlike the width and depth of research to be found in comprehension, topic familiarity in language production has been much less dealt with in the past. It has, though, become a topic of interest in recent years (most of the relevant works were published at around or after 2000). Unfortunately, the overwhelming majority of research discussed above has been about L1 spoken discourse, which was done mostly by psychologists whose interests differed from applied linguists and SLA researchers. Though we may gain some insights from what is available in L1 research, there is certainly an urgent need to explore the effects of topic familiarity in L2 oral production.

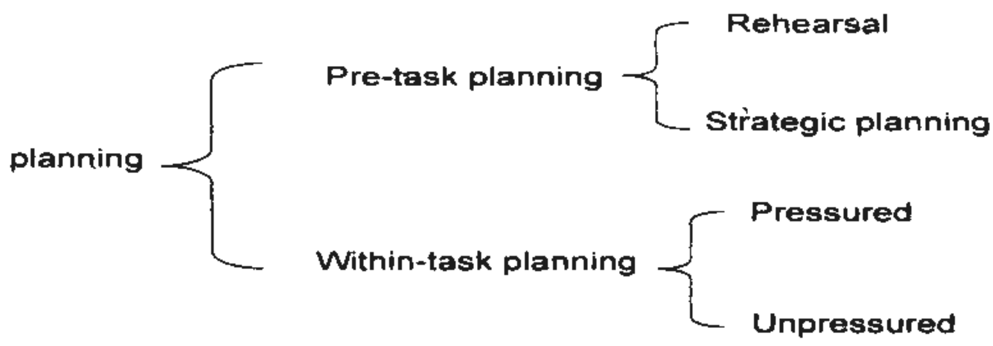
2.2.2 Planning in general and strategic planning in particular

This section first of all provides a brief introduction to the research in planning, and then goes into details about strategic planning which is the type of planning directly relevant to the present study.

2.2.2.1 An introduction to planning

All language use, either written or spoken, involves planning because one has to decide what to express and how to express it (Ellis, 2005). Planning can happen at discoursal, sentential, and constituent levels, and may take place before or during a task, resulting in different types of planning that will be talked about below. As Ellis (in press) points out, research into the effect of planning has theoretical interest because "it serves to test claims regarding the nature of variability in learner language and the validity of models of L2 speaking such as that of Levelt (1989)". In addition, studies on planning may also be of practical interests as it may have pedagogical implications that can inform teachers in task-based instruction, "where one of the options available for implementing tasks concerns whether or not to allow students time to plan and, if so, what kind of.

Figure 1.4 Types of task-based planning (Ellis, 2005, p. 4)



planning and for what length of time" (Ellis, in press)

Figure 1.4 illustrates two major types of planning by Ellis (2005): pre-task and within-task planning, depending on when the planning takes place. Each type subsumes two sub-categories. Ellis (in press) slightly revised the categorization and distinguished between three major types: rehearsal (e.g., Bygate, 1996, 2001; Lynch and McLean, 2000, 2001; Bygate and Samuda, 2005), pre-task planning (e.g., Crookes, 1989; Foster and Skehan, 1996; Skehan and Foster, 1999; Wigglesworth, 1997), and within-task planning (Yuan and Ellis, 2003; Ellis and Yuan, 2005). More recently, Wang (2009) was able to include five different types of planning in one study. The findings confirmed the effects of strategic planning in general, but on-line planning alone was found to be insufficient to make a difference in task performance unless some pre-task preparation is involved. Out of her five planning types, task repetition (cf., Bygate 1996, 2001) appeared to be the most powerful in that repeating a task achieved comprehensive effects in almost all performance areas: fluency, accuracy, complexity and lexis, whereas other types of planning can, due to the trade-off effects, only push learners for improvement in certain aspects. The following discussion will be devoted mainly to pre-task strategic planning, given its relevance to this study.

2.2.2.2 Different lengths of strategic planning

Strategic planning is operationalized by providing learners time to prepare prior to a task. Researchers have varied in the time allocated to participants. Most relevant studies have taken 10 minutes as the standard but there were some exceptions. Wigglesworth (1997) gave 107 ESL adult only 1 minute to plan as it was in a testing situation. She found that planning did not produce significant differences between the planned and the unplanned performances when they were rated by two trained raters who employed an analytic rating scale to measure fluency, grammar and intelligibility. However, the twenty-eight participants' performances were analyzed in term of a host of 'hard criteria' – fluency, complexity and accuracy measures similar to other studies. In such cases, there were significant gains in terms of fluency, complexity and accuracy for planners, especially those at a higher proficiency, and in tasks with higher cognitive load. Tavakoli and Skehan (2005) allowed 5 minutes planning time. They reported highly beneficial results on all aspects of task performance, including accuracy (in addition to the more predictable results in complexity and fluency).

Elder and Iwashita (2005) adopted a 3-minute approach in investigating strategic planning, in which very different results were obtained as both subjective holistic rating scores and objective analytic measures failed to tell the differences between planners and non-planners. They provided a long list of explanations to account for such a finding. Firstly, this simple narrative was monologic in nature and was conducted in a language laboratory. Secondly, the task instruction did not guide learners towards any focus on form. Thirdly, participants were not familiarity with speaking under planned conditions. Fourthly, three minutes could be too short, and so on (see Elder and Iwashita, 2005, for other explanations). To the present author's knowledge, only one study (Mehnert, 1998) explored the effects of different lengths (1, 5 and 10 minutes) of planning time systematically. She discovered that fluency increased as more planning time was given, but the difference between 5 and 10 minutes planning was smaller than that between 1 and 5 minutes, showing an asymptotic route. The same study found that only the 10 minutes condition produced significant greater complexity, while accuracy was able to make a difference only between non-planners and the 1 minute planners. Except for the first minute, it appeared that more planning time did not bring about greater accuracy. Except for the studies just mentioned, 10 minutes seem to be a standard planning time in most other studies which will be discusses below in terms of other categories.

2.2.2.3 Interactivity in strategic planning

Pre-task planning can be carried out in different forms according to its interactivity. For example, Foster and Skehan (1999) had three planning groups: solitary, teacher-led and group planning doing a decision-making task (a debate on what kind of person could stay on an overloaded balloon stead of being thrown out). Significant accuracy effects were observed in the teacher-led group, while greater complexity and greater fluency were achieved when participants planned on their own. Group-based planning did not lead to performance significantly different from the control group.

2.2.2.4 Attention manipulation in strategic planning

Effort has also been made to manipulate learners' attention to different areas of performance. Foster and Skehan (1999) tried to direct learners attention to different foci (towards language or towards content), but there was little evidence that these different foci had any effect on performance. In Sanguran's (2005) study, three types of guided planning – form-focused planning, meaning-focused planning and form-/meaning-focused planning – were examined. Though strategic planning in general had beneficial effects on complexity, fluency and accuracy, the three guided planning types, did not make a difference because learners tended to focus on meaning expression even if they were instructed to pay attention to other areas. Mochizuki and Ortega (2008) asked their guided planners to focus on the use of a single grammatical structure (English relative clauses). They reported a trend for non-guided planners to be more fluent than non-planners and guided planners. The guided planners did produce more accurate relative clauses than the unguided planners, but there were no group differences in the more general accuracy and complexity measures. Summing up, the above research suggests that learners do not respond very actively to guidance in planning.

2.2.2.5 Qualitative research into planning

Most of the studies on planning are based on a quantitative paradigm, but there are a few exceptions which investigated what learners actually do when they are given planning time. Wendel (1997) conducted an interview with the participants immediately after the completion of the tasks. Not too surprisingly, they responded differently when asked what they were doing in planning, but all reported that they had focused on sequencing the narrative events in chronological order. Wendel's conclusion was that learners do not benefit from planning the details of grammatical usage off-line. An introspective interview approach was employed by Ortega (1999) for the same purpose. Ortega reported that learners worked on the main ideas and organization first and then on the details, showing an identifiable manner. Quite differently from Wendel (1997), she found that learners also attend to form when planning, though considerable individual variation exists in this respect. Ortega (2005) further examined this issue by eliciting metacognitive responses from learners doing strategic planning. Using interviews again, she confirmed her (1999) study that learners did attend to form during planning. She concluded that "pre-task planning created the mental space for learners to negotiate with themselves many aspects of the language and allowed them to utilize various funds of explicit knowledge that guided their conscious attention towards areas in which they were well aware of holes and gaps visa-vis the specific task demands" (p.105-106).

The latest development in qualitative research into the actual planning processes is Pang and Skehan (in preparation) in which several features are noteworthy. First, learner self-report was triangulated by their actual performance. Second, qualitative data were related to Levelt's (1989) speaking model. Thirdly, and for the first time in these types of research, proficiency was considered seriously. They found that the cognitive processes identified during planning fit in Levelt's (1989, 1993) models well.

2.2.2.6 Theorization in planning research: Robinson Vs Skehan

Though previous research (e.g., Crookes, 1989; Foster and Skehan, 1996, 1999; Wigglesworth, 1997; Mehnert, 1998) has demonstrated that pre-task planning generally emerges as a beneficial means for improving L2 speakers' oral speech, and its pedagogical value in task-based learning is accordingly worth exploring, divergent opinions exist as in *which areas* and in *what way* planning exerts its impact on task performance. In addition to what has been discussed previously, Crookes (1989), for example, reported that planning gave rise to complexity and fluency, with accuracy unaffected. However, Foster and Skehan (1996) found that planning did promote accuracy, especially for the unguided planners, in that they probably used the planning time to rehearse language, hence greater accuracy. As Ellis (2005) mentioned recently, two positions appeared in theorizing the effects of planning on task-based performance, known as the Skehan-Robinson Debate. Obviously theorization of this kind should belong to the theoretical ground above, but the fact that both theories has developed from or has been tested against the empirical backdrop discussed above makes it a more natural place for a mention of this debate here – It follows logically from the planning research literature.

Skehan (1996a, 1998) argued for a limited processing capacity model in which L2 speakers have to trade off their aspects of language due to a limited processing capacity. L2 learners vary in the extent to which they prioritize fluency, complexity and accuracy. Some task conditions and task characteristics predispose learners to focus on fluency, others on accuracy, and yet others on complexity. Skehan (1998) drew a distinction between an exemplar-based system and a rule-based system. The exemplar system is connected with fluency as it helps learners access their memory-based system for ready-made chunks; the rule-based system, however, is associated with accuracy and complexity because the two aspects of performance require syntactic processing. Deriving from this view, planning can free up attentional resources and helps to improve performance in some areas, depending on which areas learners choose to emphasize. In general, Skehan's model may predict that strategic planning allows learners to attend to the rule-based system and become less reliant on the exemplar-based system, thus enabling them to be more willing to task risks (higher complexity) at the expense of accuracy (trade-off).

From a different perspective, Robinson (1995, 2001a, 2001b) holds a multiple-resource view of processing, arguing that L2 learners, like native speakers, are capable of parallel processing and attend to more than one aspect of performance at the same time. He believes that complexity of language is decided by the complexity of tasks, and that there is no competition between complexity and accuracy. Robinson suggests two categories of features in determining task complexity: resource directing and resource depleting. Tasks requiring higher cognitive load, like those with reasoning, are resource directing; whereas tasks given favorable conditions, for example, providing planning time, are resource depleting. Therefore the prediction which Robison makes is that planning prior to a task results in exhaustion of attentional resource, thus leading to increased fluency but decreased complexity and accuracy. In the literature there is clearly more evidence (e.g., Foster and Skehan, 1996, 1999; Mehnert, 1998; Ellis, 1987; VanPatten, 1990; Wang, 2009) in support of Skehan's model, which Robinson (2001b) conceded. That being said, given the range of interlanguage measures and the diverse operationalizations of planning across the studies, it is not easy to offer a clear-cut evaluation on the effects of pre-task planning (Ortega, 1999). A reexamination of the construct of planning would seem beneficial in future studies.

2.2.3 Proficiency

This section will relate the effects of proficiency to planning in task research, and to topic familiarity in reading and listening studies. The age-old issue of proficiency measurement will also be discussed.

2.2.3.1 Proficiency in task research

Participants in most studies in the task literature have primarily been of lower intermediate to intermediate levels for at least two reasons: First, intermediate learners are representative of general L2 learners, making the study generalizable to the largest group of students possible. Second, it is usually easier and more convenient to find participants of this kind. That said, it is obvious that loose control of proficiency cannot help us understand the potential interaction of proficiency with other variables under research. There is a need to bring in research participants with a range of proficiency levels in TBLT research, especially those with higher proficiency levels (Skehan 2009).

Speaking from a psycholinguistic perspective, low proficiency, which means limited language resources available, coupled with the fact that the three stages (Levelt, 1989) in L2 speech become controlled processes (Ellis, 2005), leads to difficulties in L2 speech production. The task literature (especially that of planning), unfortunately, does not offer much evidence regarding an interaction between proficiency and strategic planning (Kawauchi, 2005). A study by Wigglesworth (1997) was one of the few studies that considered proficiency seriously. In this study, the opportunity for one-minute planning time only raised complexity and accuracy among the high-proficiency learners, which was more evident in the case of tasks with a high cognitive load such as the picture description task. Learners of low proficiency did not appear to benefit from planning time. However, Mehnert (1998) showed that different lengths of planning time (1, 5, and 10 minutes) brought about different effects. Therefore, we may want to raise another question: will planning time also makes a different among low-proficiency learners when it is prolonged to a period like 5 or 10 minutes? Tavakoli and Skehan (2005) found that the effects of planning were strong at both high and low proficiency levels, but the 'high' proficiency learners in this study were probably low-intermediate to intermediate compared to participants in Wigglesworth (1997). Also, Kormos (2006) suggested that global self-repair behavior was not affected by different levels of proficiency.

2.2.3.2 Proficiency and topic familiarity

The picture on interactions between proficiency and the effects of topic familiarity on comprehension is not very clear. Some researchers argue that schemata are more important than proficiency in comprehension. For instance, Hudson's (1982) finding was that schematic knowledge played a more important role than proficiency in reading comprehension. Therefore lower proficient learners could also understand the passage well if they were familiar with the topic. Schmidt-Rinehart (1994) indicated no interaction between proficiency levels and topic familiarity. In this listening comprehension study, all subjects scored higher on the more familiar passage regardless of their proficiency levels. The conclusion they would draw is: topic familiarity overrides proficiency.

Other studies produced contradictory results. As mentioned earlier, Carrell (1983) reported that background knowledge only helped native speakers but not L2 learners in reading comprehension, and her explanation was that the L2 participants in this study had not yet pass the linguistic threshold to utilize topic familiarity in reading. Phillips (1990) found that prior knowledge was insignificant when learners lacked adequate reading proficiency. Douglas (2000) even pushed this argument much further by saying that lower proficiency learners are incapable of using background knowledge even if it is present, while high proficiency learners don't need background knowledge because their proficiency level can make up for the gap in this respect. This is in agreement with Chern (1993) in which proficiency appeared to be a stronger predictive variable than topic familiarity in reading retention and vocabulary gains. Therefore, we may also say, based upon these studies, that proficiency overrides topic familiarity.

2.2.3.3 Proficiency measurement

The inconsistency in the effects of L2 proficiency could be attributed to the definition of proficiency *per se*. The concept of language proficiency is never a simple construct as it relates to language competence, metalinguistic awareness, and the ability to speak, listen, read and write the language in contextually appropriate ways (Lee and Schallert, 1997). Established public proficiency tests like TOEFL, IELTS, and the Michigan Test are usually accepted as a valid and reliable way of operationalizing the proficiency construct. Nonetheless, researchers in the field are still faced with the situation that no one single universal test available, which results in heterogeneity in cross-study comparisons that is actually part of the reason for the dispute. In North America, Trammell (1991) represents one attempt to solve the problem as he suggested extending the American Council for the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines to form the basis of testing L2 proficiency for reading-for-research. In Europe, the Diagnostic Language Assessment System (DIALANG) is taken as an effective and widely recognized measure of proficiency primarily for European citizens to assess their language abilities in adherence to Europe's Common European Framework of Reference (CEFR) in 14 European Languages. Despite all this effort, attempts in this regard often end up as suggestions only, because variables like learner background and social contexts are all resistant to the use of a highly uniform proficiency test. That said, people begin to deal with possibility of equating the numeric score with a certain band score in studies like TOEFL and IELTS, which could shed some light on cross-study comparisons.

The imprecise cut-off line for different proficiency levels undermines a sound overview of its effects. What the first camp (those who argue for familiarity over proficiency) regard as 'low' could probably be 'intermediate' or even 'high' in the second. If there were a precise universal proficiency measurement, those regarded as 'low' by the second camp would only be absolute beginners, while those termed as 'low' in the first camp and those 'intermediate' to 'high' in the second are really intermediate learners. And those 'high' in the first camp are the real 'high' people. Therefore, there seems to be a linguistic threshold before learners are able to capitalize on their schematic knowledge. This hypothesis has important implication for my present study in that potential participants should be learners of intermediate or higher levels who have most probably passed the linguistic threshold for effective schematic knowledge to function.

2.3 Issues in performance measurement

Task research mainly explores three performance areas: complexity, accuracy and fluency, referred to as CAF. *Fluency* is the ongoing speech without undue pausing or hesitation (Ellis and Barkhuizen, 2005). More fluent language happens when meaning is prioritized over form in the process of task completion (Skehan 1998; Tavokoli and Skehan, 2005). Factor analyses (Skehan and Foster 1997a;

Tavokoli and Skehan 2008; Mehnert, 1998) generally confirmed two relatively independent types of fluency: breakdown fluency and repair fluency, each with a range of variables subsumed under it. Breakdown fluency is concerned with filled and unfilled pauses, so this category usually subsumes measures such as number of mid-clause pauses (Foster and Skehan, 1996), speech rate (Tavakoli and Skehan, 2005), and mean length of run (Skehan and Foster, 2005). Repair fluency is usually measured through reformulation, false starts, replacements and repetitions (Foster and Skehan, 1996; Tavakoli and Skehan, 2005).

Accuracy is the extent to which learners produce the target language in relation to the rule system of that target language (Skehan, 1996). Higher accuracy is the result of a more conservative stance on the part of the speaker to avoid errors. Two main types of measures were employed in different studies: specific and general. The specific measures focus on one particular error type, such as correct verb forms (Ellis and Yuan, 2005), proportion of correct past-tense use (Ellis, 1987; Kawauchi, 2005) or article use (Storch, 1999). The general measures generate one value to represent all errors. The classical and most widely used measure is the ratio of error-free clauses to all clauses. Skehan and Foster (2005) pinpointed a methodological flaw in such a measure in that higher accuracy is easier to achieve in shorter clauses. They proposed a modified version of accuracy measure that takes clause length into consideration. First the proportion of correct three-word clauses, four-word clauses and so on is calculated. Then the length of a clause with a certain percentage (they tried 50%, 60%, and 70%) that is correct is set as the cut-off point beyond which the participant cannot produce correct clauses at this level.

Table 2.1 Error-free clauses and clause length: Hypothetical examples (Adapted from Skehan and Foster, 2005)

	Percentage accuracy scores		
Clause length	Learner 1	Learner 2	Learner 3
2	80	80	80
3	70	80	40
4	70	40	80
5	70	70	40
6	50	70	60
7	40	60	40
8	30	40	80
9	30	40	40

Table 2.1 shows three hypothetical learner performances, with 70% as the example baseline. The first one is straightforward and he receives an accuracy score as 5. The second, though complicated by a percentage lower than 70% at the 4-word clause level, meets the criteria in the main. Hence learners two gets a score of 6. Learners three represents the case of non-successive level of performance in which a conservative criterion is adopted. S/he would be given a score of 2, the lowest level where the criterion is met.

However, these two indices of accuracy view a clause as the error count unit instead of the actual number of errors. One can argue that a clause with only one error is less serious than one with more errors, but the previous two measures will count both of them as the same. Menhert (1998) and Sanguran (2005) used 'errors per one hundred words' as another measure of accuracy, which includes all the errors that occur.

Complexity refers to how elaborated the language produced by learners is, which reflects the extent to which learners take risks (Skehan 2001). Traditionally complexity is measured by the ratio of subordination clauses to a speech segmentation unit (T-unit, C-unit and more recently, AS unit). Norris and Ortega (in press) argued that this measure works better with learners at lower but not higher proficiency levels. They proposed the length of clauses (in words) as a better index of complexity for more advanced learners. Participants in this study are intermediate to high proficiency university students who would be an appropriate group to explore the fitness of Norris and Ortega's proposal.

If fluency is more linked to an emphasis on meaning during speaking, then complexity and accuracy are more concerned with attention to the organizational and structural nature of language (Skehan 1998). As Skehan (2009) pointed out, however, the *lexical* aspects of task performance have been largely ignored in previous task research. Several lexical indices are available in the literature. Ure (1971) used lexical density in her trailblazing corpus work and showed that this is an important dimension in differentiating spoken and written language. Written language tends to enjoy a higher ratio of content words than spoken language. Halliday and Martin (1993) developed this point and defined lexical density as "a measure of the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure" (p.76). Halliday measured lexical density as the number of lexical words per clause, which is a bit different from the commonly used 'ratio of content words to total words'. A second choice to investigate lexical richness is lexical diversity, measured though the type-token ratio. The raw type-token ratio is well-known to be unreliable because of its sensitivity to text length. The longer the text is, the lower the type-token ratio will be. Different alternatives have been developed to address this problem. A widely used measure is the value of *D*, which is a corrected type-token ratio and is regarded as an indicator of the extent to which speakers draw on a larger vocabulary and return less to the same set of words (Malvern and Richards, 2002). Yet another lexical measure is available: lexical sophistication (Read, 2000), as indexed by the value of *Lambda*, which is a measure of the degree to which learners utilize more rare words (Meara and Bell 2001; Skehan 2009). Meara and Bell devised a computer program, P-Lex, to mathematically model the distribution of rare words in a text, and this generates a *Lambda* value. The higher the value, the more rare words are employed.

In addition to fluency, complexity, accuracy, and lexis, there are other possible aspects that we can look into. In TaskProfile (Skehan, 2009), a computer program used to tally task research coding results, two new interesting measures are adopted: the F-score proposed by Heylighen and Dewaele (1999) and the DB-score based on the 'involved versus information procedure' distinction in

Biber, Conrad and Rappen (1998). 'Formality' was defined by Heylighen and Dewaele (1999) as avoidance of ambiguity by minimizing the context-dependence and fuzziness of expressions. They introduced the F-score, an index based on the frequency of different word classes, as an empirical measure for formality. Nouns, adjectives, articles and prepositions are more frequent in formal styles, whereas pronouns, adverbs, verbs and interjections are more common in informal styles (see table 3.7 in Chapter 3 for the formula). Heylighen and Dewaele (1999) reported that the factor analyses conducted using available corpora in 7 languages all confirmed a similar pattern roughly equal to the F-score. Biber, Conrad and Rappen (1998) proposed a 'involved versus informational production' to distinguish between personal and formal dimension of language use. The 'involved' styles includes higher proportion of private verbs, that-deletion, contractions, present tense verbs, second person pronouns, and so on (see Biber, Conrad, and Reppen, 1998, p.148, for the complete list), while the 'informational' style is just the opposite. The F-score and the DB-score operationalize the construct of formality differently (though Biber, Conrad and Reppen did not explicitly mention 'formality'), but there should be considerable common ground shared between them. So far there has been no study comparing the two measures, but a certain degree of correlation can be predicted to exist.

Chapter 3 Methodology

This chapter provides a detailed account of the participants, the measurement of their English proficiency as the criteria for grouping, the tasks, the study design, the actual procedure of data collection, the performance measures (dependent variables), and the coding scheme. Description in the chapter, though retrospective in nature, is based on the research logs kept throughout the study.

3.1 Participants

80 undergraduate students aged between 18 and 24 from the Chinese University of Hong Kong volunteered to participate in this study. They were selected from a larger pool of students (102) recruited through the Campus Mass Mailings service offered by the ITSC Department. Among them, 40 students were computer science majors with the other 40 being medical science majors. There were 50 female students (Medicine: 35 and Computer: 15) and 30 male students (Medicine: 5 and Computer: 25).

As for their linguistic background, all the students were native Chinese speakers (77 with Cantonese and 3 with Mandarin Chinese as their mother tongue). Except the 3 mainlanders (all computer science majors), all final cadidates were raised and educated in Hong Kong. Candidates with overseas living or education experience were excluded from the final group of participants. English was their common L2. They had learned English for 12 - 16 years by the time they took part in this study. While English was the sole medium of instruction in the classes for medicine majors, Cantonese (and at times code-mixing with English) was used in approximately one fourth to on third of the classes for the computer students. For both majors, only English textbcɔks were employed in the courses.

All the students received a small honorarium (50 HK dollars) upon the completion of the tasks. Table 3.1 provides a detailed description of their background information.

Table 3.1 Background information of the participants

Major	Computer			Medicine	
	40			40	
Gender	Male			Female	
	30			50	
Study yea	Year 1	Year 2	Year 3	Year 4	Year 5
	25	30	13	10	2
Age	18-19	20-21	22-23	24	
	42	27	10	1	
L1	Cantonese			Mandarin	
	77			3	
L2	English for all				

3.2 Proficiency test

The proficiency test employed in this study was a C-test borrowed from Dornyei and Katona (1992) (see Appendix 1). A C-test is similar to a cloze test as both of them are based on the same theory of closure or reduced redundancy (Alderson, 2000). In C-tests, the first and last sentence remain intact, but the second half of every second word in the rest of the text is deleted and has to be restored by the test-taker. Though the C-test has received criticism from Alderson (2000), most studies show that the format is a reliable measure of general proficiency. Dornyei and Katona (1992) found that the C-test is *reliable* (the internal consistency coefficients are very consistent, .75 and .77 respectively for university English majors and secondary students) and *valid* (the C-test is significantly and highly correlated with 4 different general proficiency tests such as the TOEIC). Cronbach's *alpha* reached .84 in Daller and Phelan (2006). Klein-Braley and Raatz (1984), Cohen, Segal and Bar-Siman-Tov (1984), Klein-Braley (1985, 1997), and Grotjahn (1995) generally supported such a claim with *written* tasks. More importantly here, the C-test was reported to be highly correlated with *oral* tasks as well in recent studies (e.g., $r = .64$ in Arras, Eckes and Grotjahn, 2002, and also in oral lexical performance in Daller and Xue, 2007).

In a Chinese context, Dornyei and Katona's (1992) C-test was piloted among students from the same university as the present participants in 2008, with encouraging results (see section 4.2 in Chapter 4 for details). More recently, Dai

(2009) reported the same test used among Chinese students with quite a high internal reliability (Cronbach's *alpha* = .770) and good concurrent validity with China's nation-wide English proficiency test – the Chinese College English Test (CET-4), reaching a correlation coefficient r = .633 (p < .01). Specifically for the present study, and similar to Dai's (2009), the internal reliability was high (Cronbach's *alpha* = .740). Though the concurrent validity of this C-test with a general proficiency test in Hong Kong could not be tested because participants in the main study entered university with various English exam results, such as A-Level UE, HKCEE, TOEFL, and IELTS, the literature over the years and the pilot studies have lent support to the use of this C-test as a good tool for the prediction of learner's proficiency level.

3.3 Tasks

As concluded in the literature review (subsection 2.2.1.2), an essential step to effectively investigate the effects of topic familiarity is to ensure the validity of the topics on which learners are to perform a task. The extent to which learners are familiar with different topics should be clear-cut to rule out the possibility that learners possess similar background knowledge for all subject areas under investigation. The most common practice in the literature for such a purpose is to administer a post-task survey eliciting ratings through questionnaires with Likert-scale measurement. However, this did not appear to be always reliable (e.g., Hammadou, 1991). In addition to the subjective judgment from participants, researchers need to impose some 'hard' criteria (such as matching or mismatching the topics with one's native culture, as Lee (2007) did). An additional problem lies in the comparability between the topics. All topics should be as comparable as possible except for learners' background knowledge about them.

To address the first problem, the present study imposed 'double insurance' to make sure that no participants have similar background knowledge for both topics. Firstly, topic choices were determined on the basis of match or mismatch of a participant's academic discipline in the university. The natural virus topic (see Table 3.2 below), for example, is a familiar topic for medicine majors but an unfamiliar topic for computer majors. This constituted an objective criterion. Secondly, after each task, the participant was required to rate his/her familiarity with the topic (see Appendix 3 for the survey form), which was apparently a more subjective criterion. Performances by candidates who indicated in the survey that they had had similar background knowledge about both topics or had been more familiar with the topics mismatching their majors were all excluded from the final data processing.

To tackle the second problem, that of comparability, two solutions were also pursued. First of all, the two topics were intentionally designed to be as equal to each other as possible in terms of coverage of content areas, in schematic structure as well as the staging of the steps in the description. Second, as will be explained in section 3.4 (Study design), each cell contained 10 medicine majors and 10 computer majors to counterbalance the topic effect.

Another advantage of the two topics is that, though one is likely to be more familiar with the topic matching his/her specialty, no candidates would be intimidated or kept from speaking by the topics. It was assumed that a university student in Hong Kong should have already learned some knowledge of biology in their secondary school and had experience of catching a cold due to a virus. More-

Table 3.2 Task topics

Communicative context	You are a specialist in the field giving a presentation to a group of university students who are neither medicine nor computer majors but are interested in the topics.
Topic 1	Please describe in detail the general process of the infection of virus in a human body, the possible consequences, and the general procedure for dealing with a virus-infected person.
Topic 2	Please describe in detail the general process of the infection of virus in a computer, the possible consequences, and the general procedure for dealing with a virus-infected computer.

over, the computer and the Internet have become an indispensible part of life for students, and computer viruses should not come as an alien topic to the participants. This assumption was confirmed in pilot study 3 (See section 4.3 in Chapter 4 for details).

These two topics are shown in table 3.2 above.

The communicative context was orally explained to the participants by the researcher in Cantonese (or Mandarin in three cases). The topics were presented separately in each task in both English and Chinese, written on a card to ensure correct understanding. The relationship between the academic backgrounds and the topics is shown in the Table 3.3.

Table 3.3 Topic design

	Topic 1	Topic 2
Medicine Majors	+ familiar	– familiar
Computer Majors	– familiar	+ familiar

Topic familiarity served as a within-subject independent variable in the present study. That is, all students performed two speaking tasks, one of which was on a familiar topic and the other, unfamiliar. The order of topics was counter-balanced among participants to avoid the interference of practice effects. Each of the two speaking tasks (familiar and unfamiliar) was performed under two planning conditions. The 40 non-planning participants had to start speaking once they were told the topics. The 40 strategic planners were given 10 minutes to plan on their own (see section 3.5 for a detail description of the procedures).

3.4 Study design

The present study constitutes a 2 × 2 × 2 split-plot factorial design (cf., Gardner, 2001, p.127-153). There were two between-subject variables (planning and proficiency) and one within-subject variable (topic familiarity), with each being a two-level variable, as shown in Table 3.4 below. Basically, the 80 candidates were evenly divided into a planning group and a non-planning (control) group. Within each group, there were two subgroups, each containing 20 high and 20 intermediate proficiency learners, based on their C-test results. These 20 candidates consisted of 10 computer majors and 10 medicine majors to counter-balance any topic effect. That is, when there is a topic familiarity effect, we can be more confident that it is not simply because one topic is easier than the other, since each cell performs exactly the same topics.

Given the fact that disciplines were not regarded as an independent variable in this study, the sample size in each cell therefore reached 20. Because every participant performed two tasks, the 80 candidates produced 160 data points in total.

Table 3.4 Breakdown of the study design (Disciplines are italicized to indicate that they were not regarded as an independent variable in this study)

Planning conditions	Proficiency	*Disciplines*	Sample size	Tasks
Planners	High	*Computer*	10	Familiar
				Unfamiliar
		Medicine	10	Familiar
				Unfamiliar
	Intermediate	*Computer*	10	Familiar
				Unfamiliar
		Medicine	10	Familiar
				Unfamiliar
Non-planners	High	*Computer*	10	Familiar
				Unfamiliar
		Medicine	10	Familiar
				Unfamiliar
	Intermediate	*Computer*	10	Familiar
				Unfamiliar
		Medicine	10	Familiar
				Unfamiliar

Such a design requires proficiency control over all grouping variables. Table 3.5 shows such a control over a range of variables that might induce interference to the robustness of the design, using the Univariate procedure in SPSS 17 because only one dependent variable (C-test scores as proficiency) was involved. Specifically, there is no significant difference in proficiency between male and female students, or between planners and non-planners, or between those who did the familiar tasks first and those who did the unfamiliar tasks first, but there is a very significant difference between the High and the Intermediate learners. No interaction effect is found at any level. The above results confirmed the robustness of the grouping except for one variable – discipline.

Table 3.5 Proficiency control for each grouping variable

Source	Type III Sum of Sq.	df	Mean Square	F	Sig.
Gender	6.43	1	6.43	.23	.635
Planning	12.92	1	12.92	.46	.502
Discipline	904.28	1	904.28	32.03	.000
Counterbalancing	61.92	1	61.92	2.19	.145
Proficiency	2993.01	1	2993.02	106.01	.000

In an ideal situation, no proficiency difference should exist between computer majors and medicine majors. Otherwise, this study may run the risk of the discipline × proficiency interaction effects because of the internal proficiency discrepancy in each cell (e.g., the computer High is lower than the medicine Intermediate). In order to probe into this problem, a one-way ANOVA was performed to test the difference between computer High, medicine High, computer Intermediate, and medicine Intermediate groups. As indicated by Table 3.6, even though the four cells are all different to one another, the High groups are always higher than the Intermediate (i.e., the computer High is higher than the medicine Intermediate). This more reassuring information suggests that the proficiency effects, if any, cannot be attributed simply to the discipline interference.

Table 3.6 Proficiency means with an S-N-K post hoc test

			Subset for alpha = 0.05			
	Proficiency	N	1	2	3	4
Student-Newman-Keuls[1]	Com_inter	20	30.95			
	Med_Inter	20		39.45		
	Com_high	20			44.55	
	Med_high	20				52.80

A more serious aspect of this issue is that, if the two discipline groups perform differently, is it because of their academic background or their proficiency? The best way out is to hold the proficiency level constant and see if the two discipline groups differ in performance. If not, we can then be confident that any differences between the two groups are due to their proficiency, not academic specialties, and the original design can be kept intact. If yes, we then know academic background *per se* makes a difference and should be taken as an independent variable. In such a circumstance, the 10 medicine majors and 10 computer majors in the original cell should be treated separately. Given that this solution involves participants' performance, we delay the discussion in Section 5.1 until the next chapter (Results).

3.5 Procedures

This section is a step-by-step description of the data collection process. Before the actual task administration, all participants filled out a background information form (Appendix 2) and took the C-test (Appendix 1), based on which they were grouped according to the study design presented in table 3.4 above.

Step 1: Before the main tasks, a 5-min training session was conducted for two purposes. Firstly, the researcher explained the general process, telling participants they would perform two speaking tasks. Any questions could be raised at this point to be clarified by the researcher. Secondly, an MP3 recorder was placed before the participant during the training period so that they could get used to it.

Step 2: The researcher placed a card with one of the two topics on it before the participants and immediately started reading it. Participants in the non-planning group were asked to start speaking once the researcher finished reading. The strategic planning group was given 10 minutes (Crookes, 1989; Foster and Skehan, 1996; Wang, 2009) to plan on their own during which time they were instructed to take notes so that they could be mentally engaged. All materials, including the paper and pens were provided by the researcher. The notes were removed when the planners began speaking, as the participants had been informed during the training session. However, the prompt remained on the desk until they completed the task.

Step 3: After the speaking tasks, the participants were asked to rate their degree of familiarity with each topic by filling out the familiarity survey form (Appendix 3). Participants who indicated equal familiarity with both topics, or who had higher scores for the supposedly unfamiliar topics, were excluded from the study.

3.6 Performance measures

The present study followed the research tradition reviewed in section 2.3. to investigate fluency, complexity, accuracy and lexis, with the two formality measures added as well. Table 3.7 below provides a brief description of these dependent variables.

The list below further operationalizes the repair fluency terms mentioned in Table 3.7, following Foster and Skehan (1999).

- *Reformulations:* Either phrases or clauses that are repeated with some modification to syntax, morphology, or word order.
- *False starts:* Utterances that are abandoned before completion and that may or may not be followed by a reformulation.
- *Repetitions:* Words, phrases or clauses that are repeated with no modification whatsoever to syntax, morphology, or word order.
- *Replacements:* Lexical items that are immediately substituted for another.

Table 3.7 Dependent variables

General Category	Variable Name	Description	Sample Studies
Fluency	Pausing	The number of pauses and the amount of silence per 100 words. In the present study it is operationalized as any break of 0.4 second or longer.	Foster and Skehan (1996)
	Repair Fluency	This measure is distinct from breakdown fluency and should be treated separtely. In the present study it is operationalized as the total number of repetitions, replacements, false starts and reformulations per 100 words.	Foster and Skehan (1996)
	Speech Rate	A pruned speech rate is investigated here because it shows the 'real' speed of the speaker. It is operationalized as the total words per minute after deletion of filled pauses, reformulations, replacements, false starts, and repetitions.	Tavakoli and Skehan (2005)
	Mean Length of Run	The number of words uttered before any breakdown or repair is encountered.	Skehan and Foster (2005)
	Phonation time	The ratio of voicing time to the total time of utterance.	Tavakoli and Skehan (2005)
Accuracy	Error-free Ratio	The ratio of error-free clauses to all clauses.	Foster and Skehan (1996)
	Errors per 100 Words	The number of errors in every pruned one hundred words.	Mehnert (1998)
	70% Accuracy Clause length[1]	The length of a clause with 70% of all clauses of the same length correct is set as the cut-off point beyond which the participant cannot produce correct clauses at 70% level.	Skehan and Foster (2005)

[1] Following Skehan and Foster (2005), for example, if 50% of all 5-word sentences but lower than 50% of all 6-word sentence are correct, then with 50% as the threshold, the accuracy score is 5 in that L2 speech. This study calculated 50%, 60%, and 70% as the thresholds, but only the 70% value is reported in this study because it was found that 70% appeared to be a better threshold in differentiating accuracy performance among learners of higher proficiency, such as those in the present study.

General Category	Variable Name	Description	Sample Studies
Complexity	Clauses per AS unit	The ratio of subordinate clauses per AS unit.	Foster, Tonkyn, and Wigglesworth (2000)
	Words Per AS Unit	The average word number in all AS units.	Norris and Ortega, in press
	Words Per Clause	The average word number in all clauses	Norris and Ortega, in press
Lexis	Lexical variety: (*D*)	Corrected type-token ratio, an index of the extent to which the speaker avoids returning to the same set of words.	The VocD sub-routine in CLAN program by MacWhinney, (2000); Malvern and Richards, (2002)
	Lexical sophistication: (*Lambda*)	The extent to which speech contains difficult or rare words.	P-Lex program by Meara and Bell (2001), modified by Skehan (2009)
	Lexical Density	The ratio of content words to total words.	Daller, Milton and Treffers-Daller, (2007, p13)
Formality	F-score	F=(noun frequency + adj. freq. + art. freq. – pron. freq. – verb freq. – adv. freq. – interjection freq. + 100)/2	Heylighen and Dewaele (1999)
	DB-score	The proportion of private verbs, that-deletion, contractions, present tense verbs, second person pronouns, do as pro-verb, analytic negation, demonstrative pronouns, general emphatics, first person pronouns, pronoun it, be as main verb, causative subordination, discourse particles, indefinite pronouns, and so on.	Biber, Conrad and Rappen (1998, p148)

Different studies employed different measures for the same construct, as reviwed in section 2.3 (Issues in performance measurement). For instance, Foster and Skehan (1996) regarded "ratio of error-free clauses" as a measure of *accuracy*, and Mehnert (1998) took "errors per 100 words" for this construct. Rarely have we seen a study that includes all existing measures and examines their relationship. The present study aims to fill this gap and tries to involve most of the available measures in the literacture. The inclusion of all these measure enables the study to examin various measures in one study and provides a level playing groud for comparisons between them. Moreover, the comparisons might have potential methodological implications – the factor analyses in section 5.10 (Underlying constructs in L2 oral production) would shed light on which of the several measures under the same category appears to be the most significant (have the highest loadings) and therefore may become a better measure for that construct in future studies.

3.7 Data processing and coding scheme

All the speaking tasks were recorded with MP3 digital recorders and transcribed using the CHILDES transcription system. Below is an example of data coding from a high proficiency computer major performing an unfamiliar task with ten-minute planning time:

*CCT: I may say why you get virus is when immune system become weaker, the virus can infect your body easily.

%mor: pro|I aux|may v|say adv:wh|why pro|you v|get n|virus v|be&3S conj:subor|when adj|immune n|system v|become adj|weak-CP det|the n|virus aux|can v|infect pro:poss:det|your n|body adv:adj|easy-LY .

%snd: <00.30.60><00.43.16>

%CCT: er {when you get the virus} * (0.878) er I may say errfr ::: er (0.660) {when you} # (0.580) er why you get virus err_m_s :;:b is when immune system become weaker err_m_s :;:a the virus can infect your body easily . errfr ::: |

The first line (*CCT:) is a pruned line with all filled pauses and repairs excluded. It is used for the calculation of lexical diversity (*D*). If repairs or filled pauses like 'er' and 'mmm' are included, then the value of *D* is very likely to become a by-product of repair fluency instead of a lexical index. In such a circumstance, the more repairs and fillers one produced (hence less fluent), the lower the *D* value will be, as *D* reflects the extent to which one recycles the same set of words.

The second line (%mor:) is the pruned AS unit tagged for their word classes, which is the basis for the formality measure (F-score). The tagging is automatically done by two subroutines within the CLAN software (MacWhiney, 2000). The MOR subroutine first tags all possibilities (e.g., the word 'present' can be a noun or a verb). Then the POST subroutine will determine the correct word class of the word according to the context. Finally, this second line is checked manually because there are still a very small percentage of errors after the POST procedure.

The third line (%snd:) indicates the beginning and ending time of this AS unit. Together with the pauses inside brackets (e.g., (0.878)) in the fourth line, it forms the basis of tallying breakdown fluency. The timing is precisely measured by the audio editing software 'GoldWave' through the use of its audio waveform.

The fourth line (%CCT:) carries most of the codes. Fluency codes include repairs (e.g., * represents repetition, and # is a false start), filled pauses (such as 'er' and 'mmm'), and unfilled pauses (e.g., (0.660)). Err_x_x represents various error types. For example, 'err_m_s' means a minor syntactic error. And different codes like ':::' and ':;:b' represents different clause types (::: is a main (superordiante) clause and :;:b is finte subordinate clause that occurs before the main cluase). TaskProfile (Skehan, 2009) automatically recognizes all the codes and provides an output form with almost all the measures employed in this study.

'CCT' in the first and the fourth lines is a code for a specific participant whereby the research can trace the source of the transcript from the audio recording.

Appendix 5 provides 8 sample files, all coded according to the above-mentioned scheme, as representatives from the 8 cells in the study design.

3.8 Statistical procedures

A number of statistical procedures were employed in this study according to the appropriacy of their application. The major statistical procedure was a repeated measures multivariate analysis of variance (MANOVA) that applies to a study design with more than one dependent variables for both between-participant and within-participant comparisons. In addition to the MANOVA, T-tests, Pearson correlations, Factor Analyses, reliabity test, and so on were adopted for various groups of normally distributed data. In case of data that were not on a normal curve, non-parametric tests, such as Wilcoxon signed rank test and Mann-Whitney *U* test, would be applied. Non-parametric tests were mostly used in Chapter 4, Pilots Studies, due to the small sample sizes in that certain context.

In this dissertation, normality tests (both Shapiro-Wilk test and Kolmogorov -Smirnov test) were performed before other statistical procedures to ensure that the normality assumption for parametric tests could be met. There will be footnotes in later text to indicate non-normality distribution when non-parametric tests were employed. However, the default assumption of normal distribution of data, which occurred in most cases in this study, would not be specified.

Chapter 4 Pilot studies

Three pilot studies were conducted prior to the main study. The purpose of these studies was to test the validity and reliability of a C-test in a Hong Kong context, to explore the robustness and effectiveness of the task design, as well as to trial the procedures for the main study. The three pilot studies will be reported chronologically.

4.1 Pilot study 1: A self-devised C-test

The literature has, in the main, supported the use of a C-test as an effective predictive measurement of general proficiency. Following the theory of reduced redundancy and examples in the literature, a C-test was constructed by the researcher consisting of 50 blanks (Appendix 4). This C-test was piloted with 20 undergraduate students from the same university as the participants in the main study. Correlations were calculated between this C-test and two general proficiency tests: the English test in the Hong Kong Certificate of Education Examination (HKCEE) and the Use of English test in the Hong Kong Advanced Level Examination (A-Level), both conducted by Hong Kong Examinations and Assessment Authority. The HKCEE is usually sat for by students at the end of their 5 years of secondary education; the A-level is taken at the end of the 7th year for university matriculation. According to Wikipedia, the HKCEE is comparable to the British GCSE O'Level standard, and the comparison between Use of English in the A-Level and the International English Language Testing System (IELTS) is listed in table 4.1.

Table 4.1 Equation of Use of English (A-Level) and IELTS

Use of English (A-Level)	A	B	C	D	E
IELTS	7.41 – 8.30	6.92 – 7.40	6.51 – 6.91	6.03 – 6.50	5.40 – 6.02

In this study, the marks of A to E were quantified as 5 to 1, with 1 point between each letter mark. The result showed that this C-test was too easy (average: 42.8 out of 50) for a group of relatively high proficiency learners and therefore had low discriminative power (S.D. = 3.88). Also, the correlation between this C-test and the English test in HKCEE only showed a trend towards significance ($r = .417$, $p = .067$) and no significant correlations with Use of English in A-Level. Therefore, this C-test was abandoned and a revised version was sought.

4.2 Pilot study 2: A revised C-test

The revised C-test consisted of four passages with 78 blanks reported to be valid and reliable in Dornyei and Katona (1992, see Appendix 1), and was tested among 45 undergraduates from the English department as well as the education faculty of the Chinese University of Hong Kong out of whom 42 valid samples were collected. The general picture is showed in table 4.2:

Table 4.2 Descriptive statistics of the second pilot study (N=42).

	Min.	Max.	Mean	S.D.
HKCEE English	2.00	5.00	3.93	.95
A_Levl Use of English	2.00	5.00	3.43	.77
C_Test	27.00	64.00	50.21	7.40
Difficulty Rating	1.00	4.00	2.12	.67

Generally, the students reported that this C-test was difficult (5 was 'very easy' and 1 was "very difficult" in the questionnaire, and the mean was 2.12). The mean score was 50.21 out of 78, with an S.D. = 7.40, suggesting a better discriminative power than the first pilot study (S.D. = 3.88). Because the scores of the HKCEE and the A-Level were not normally distributed (K-S normality test, HKCEE $p < 0.001$, A-Level $p < 0.001$), a Spearman's RHO test (2-tailed) was performed to explore the concurrent validity of this C-test with the two established English proficiency tests. The results were encouraging because, as expected, these proficiency tests correlated very significantly with one another. HKCEE English test and Use of English in A-Level were significantly correlated ($r = .48$, $p < .001$). The r of only 0.5 can be attributed to two possible sources. First, the participants' proficiency changed during the two years' time between the two exams. Second, the possibility that the two tests measure slightly different things cannot be ruled out. This is also true with the correlations between the C-test and the HKCEE (r = .46, p < .01), and between the C-test and the A-Level (r = .41, p < .01). Given the fact the two most authoritative Hong Kong English exams correlated at r = .48, the concurrent validity of C-test with them at an r of .41 to .46 appears to be good enough for research purposes.

If the validity was acceptable, the internal reliability would also need examining. There are 4 passages in this C-test and the reliability (Cronbach's *Alpha*) was 0.64. The correlations in Table 4.3 show that passage 1 significantly correlates with passage 2, while passage 3 significantly correlates with passage 4, which seems to indicate that this C-test measured more than one single aspect of proficiency. The item-total statistics in Table 4.4 suggest that the four passages all positively contributed to the reliability of this C-test and any passage deleted would be detrimental to the overall reliability.

Table 4.3 Correlations between the four passages in the C-test

	P1	P2	P3	P4
P1	1			
P2	.503**	1		
P3	.29	.09	1	
P4	.21	.25	.501**	1

Table 4.4 Item-total statistics the C-test

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
P1	35.43	36.84	.45	.31	.554
P2	38.83	36.55	.36	.30	.607
P3	35.26	34.20	.42	.31	.571
P4	41.14	30.47	.46	.30	.543

This C-test correlated reasonably with comprehensive proficiency tests but its reliability did not reach the level that Dornyei and Katona (1992) claimed (.77 and .75 in two proficiency groups). Several reasons may exist here. Firstly, since there are a wide range of proficiency levels among students at CUHK, it may be that the English majors are at the high end. Then this C-test could not fully display its discriminative power among a group of relatively homogenous students in terms of proficiency. Secondly, the sample size was rather small here (42 people). After including all students in the main study, who are a more natural group of students, a better reliability score might be obtained (see section 3.2 above for the confirmation of such a hypothesis: Cronbach's alpha was .740 in the main study.).

In addition to the validation processes involving concurrent validity and reliability, several other important improvements were made after the two pilot studies. Firstly, A Chinese version of the instructions for reporting their previous English exam results was added to the C-test because some participants raised the question about which exam results they need to report. The revised version highlighted that participants should report the marks of the English tests, not the overall results of the HKCEE and the A-Level. Secondly, judging from the finishing time of the participants in the pilot studies, 20 minutes was adopted as the time limit for taking this C-test, which ensured that most students could finish the blanks of which they knew the answers. Thirdly, though Hong Kong students are familiar with ordinary cloze tests, none of the pilot study participants reported to have seen a cloze test of this kind where the second half of every other word is deleted. Therefore, it was felt that a brief explanation of the test format would be necessary before the actual test, so that students would not have to waste time wondering about the correct procedure. Given the encouraging validation processes and the lessons learned from the pilot study, the Dornyei and Katona (1992) version C-test was adopted as the proficiency measure in the main study.

4.3 Pilot Study 3: Testing topic choice, tasks implementation and procedures

Pilot study 3 aimed to test the topic choice and to trial task implementation and administration procedures. Seven undergraduates from CUHK, including both computer and medicine majors, were invited to participate in pilot study 3. The participants' information is summarized in Table 4.5. These learners each performed 2 tasks on the 2 topics in table 3.2, following the procedures in the main study. After the speaking tasks, they were asked to fill in a survey form (Appendix 3), rating how familiar they were with the two topics. Then the researcher held an open interview with each of the participants for their feedback on the topics, tasks and procedures. The results will be reported in the following subsections. Based on the pilot results, adjustment and improvement were made in a number of areas in the main study.

Table 4.5 Participant distribution in different planning conditions and disciplines

	computer	medicine
Planners	0	3
Non-planners	2	2

4.3.1 The robustness of the topics: Participant rating

The rationale behind the choice of these two kinds of virus is that participants will be able to say something on each topic given their encyclopedic knowledge, but a familiar topic may encourage them to go further since they are in possession of the relevant schematic knowledge in greater width and depth. The rating of the participants confirmed such a hypothesis (see the table 4.6). The lowest rating for the unfamiliar topic was 2 (= know something), which means nobody will be kept from speaking simply because they know nothing about the virus. At the same time, a Friedman Test[2] showed that there is a significant difference between their familiarity with the two types of topics ($\chi^2 = 6.00$, $P = 0.014$), suggesting that the subjective and objective criteria converge well.

Table 4.6 Familiarity ratings

Rating	N	Mean	S.D.	Minimum	Maximum
Familiar topics	7	4.00	.58	3.00	5.00
Unfamiliar topics	7	2.44	.79	2.00	4.00

4.3.2 Robustness of the topics: Participant performance

The task performances were recorded and the speech samples were transcribed. The transcription was sent back to the participant for his/her verification, especially for a set of less frequent terms related to medicine or computer science, such as 'B lymphocytes', 'phagocytosis' and 'BIOS'. At this piloting stage, only some preliminary analysis was conducted. Table 4.7 below shows some descriptive statistics on their performance:

As expected, speakers produced more total words (Wilcoxon Signed Rank Test[3], $Z = -2.03$, $p < .05$) when they encountered their own familiar topic, but this

[2] The rating scores were not normally distributed, so a non-parametric test was employed instead of a paired- sample T-test.

[3] The data is not normally distributed.

Table 4.7 Descriptive statistics of familiar and unfamiliar tasks (N=7)

	Familiarity	Min	Max	Mean	S.D.
Total words	Fam	252.0	1461.0	600.9	419.1
	Unfam	97.0	421.0	313.1	123.6
Total time (Sec.)	Fam	143.9	680.0	298.7	192.2
	Unfam	73.5	198.0	165.9	44.6
Speech rate	Fam	94.6	149.5	120.4	21.3
	Unfam	79.2	135.1	109.3	22.1
Raw type/token ratio	Fam	.24	.40	.34	.06
	Unfam	.31	.56	.40	.08

was not achieved at the cost of significantly longer time ($Z = -1.86, p = .063$). This naturally lead to a higher speech rate (Paired-sample T-test[4], $t = 3.4$, $p = .01$). As for lexis, only raw type-token ratio (TTR) was calculated at this stage. A well-known shortcoming of the raw TTR is that it inevitably decreases as the text becomes longer. Given the fact that there was no significant difference ($p = .11$) in the raw TTRs between the familiar speech (600.9 words) and the much shorter unfamiliar speech (313.1 words), we can predict that the familiar speech might in fact contain higher corrected TTRs.

As for planning, if judging for any trend is done on the basis of simple means,

[4] The data is normally distributed

the planners uttered more words (352 > 283) in a quicker fashion (114 > 105 words per minute) under the *unfamiliar* conditions. But these planners did not achieve greater fluency (119.8 ≈120.7 words per minute) and even spoke less (498 < 677 words) under the *familiar* conditions. This result was very interesting in that it suggested an interaction effect between familiarity and planning – when speaking about familiar topics, the non-planners could speak as fluently as, or even faster than, the planners, showing a strong topic familiarity compensatory effect for the unplanned condition. In general, the planning effects were less strong than the topic familiarity effects as the Mann-Whitney U test did not display any significant difference between planners and non-planners in terms of total words, speech rate or raw TTRs (Table 4.8). Another intriguing point here is that the planning conditions had no influence on the participants' rating of their familiarity. The opportunity of ten-minute planning time for learners to recall and organize their prior knowledge did not lead them to feel more familiar with the unfamiliar topics ($p < .01$), which further validates the topic choice.

Table 4.8 Descriptive statistics of planned and unplanned tasks (N=7)

	Planned	S.D.	Unplanned	S.D.	Sig.
Total words	425.33	168.01	480.75	345.69	.724
Speech rate	120.29	23.42	109.98	23.48	.724
Raw TR	.35	.03	.40	.09	.289

Note: A Mann-Whitney U test was employed because data were not normally distributed.

4.3.3 Testing the procedures of the tasks

Another important purpose of pilot study 3 was to trial the procedures and administration in its actual implementation, and to gain feedback from participants. After the two speaking tasks, the seven pilot participants were all interviewed with open or semi-open questions like "what do you think of the tasks", "will the topics be too difficult or too easy, and why?" and "(for planners) how do you feel about the 10 minutes planning time? why?". The feedback generally supported the task procedures and also provided insights to the following aspects:

Firstly, the students all reported that they had understood the general procedures (for example, to start immediate or to plan for 10 minutes). One planner complained that he did not realize that his notes would be taken away when he spoke. Instead of an outline, he wrote down a passage in full sentences which he intended to read. Therefore, it was explicitly announced to the planners in the main study that the paper for jotting down notes would be removed when speaking started. To avoid turning the solitary planning into guided planning, no instructions about the format of note-taking were offered. The participants were left to own decision as whether to plan in point form or in full sentence.

Secondly, one participant claimed that the 10-minute planning time was too long for him: "I only needed 5 to 6 minutes and the rest of the time was boring". However, other participants reported to have liked the 10-minute planning because they can "think over what I knew" or "rehearse after jotting down the main points".

To follow the general research tradition (e.g., Crookes, 1989; Foster and Skehan, 1996; Kawauchi, 2005) and enable cross-study comparisons, this 10-minute schedule remained unchanged in the main study.

Thirdly, before participants' opportunity to speak, a card with a virus topic in English was read out loud and placed in front of them, as done in the main study. However, two non-planners asked the researcher in Cantonese to confirm their translation of the topics so as 'not to misread the topics'. This results in half to one minute extra time for the non-planners, which, according to Mehnert (1998), would potentially make a difference in a number of performance areas. A Chinese version of the topics was then added to the English version on the prompt in the main study in order that the non-planners would not be turned into one-minute planners and affect the validity of the study.

In spite of these problems, participants in general felt comfortable with the tasks and procedures. The speech content indicated that they correctly understood the topics, and the processes of task implementation (training, speaking, questionnaires and interviews) were quite smooth. With some minor adjustments mentioned above, the tasks and procedures were employed in the main study.

To sum up, pilot study 1 and 2 served to develop a better version of the C-test as the proficiency measurement for grouping, while pilot study 3 validated the topic choice and tested the procedures of task implementation. These pilot studies helped to improve the quality of the main study and allowed the main study to proceed in its intended direction.

Chapter 5 Results of Main Study

This Chapter will first report on the general picture of the results on the basis of two multivariate analyses of variance (one MANCOVA and one MANOVA). Then Univariate statistics will be presented to look specifically into each individual dependent variable for a detailed account of the results. At the close of the chapter, results of two factor analyses will be discussed in the hope of gaining some insights into the underlying constructs emerging from the performance to prepare the ground for potential theorization in Chapter 6 (Discussion).

5.1 MANCOVA: resolving proficiency difference between groups

Multivariate analysis of covariance (MANCOVA) is similar to multivariate analysis of variance (MANOVA), but allows us to control for the effects of supplementary continuous independent variables, i.e., covariates. Covariates are variables which have effects on the dependent variables, but their effects are not of interest and thus should be controlled. In experimental design, covariates are usually the variables not controlled by the experimenter, but still having an effect on the dependent variables (c.f., Meyers, Gamst, and Guarino, 2006). If we regard proficiency as a covariate in this study, then we would be able to examine if the disciplines *per se* will affect the performance. This method is proposed in response to the proficiency problem raised in section 3.4 "Study design": if there is a different between the two majors, is it because of their academic background or simply because of their proficiency levels?

A range of measures (see Table 5.3) were drawn into the present MANCOVA analysis based on the principle discussed in the section 5.2. Table 5.1 detailed the results of the main effects of all independent variables that could impose potential effects on the dependent variables.

Table 5.1 MANCOVA results of performance (main effects)

Effects	Pillai's Value	F	BGdf	WGdf	Sig.
C-test (proficiency)	.372	5.82	6	59	.000
Gender	.017	.17	6	59	.984
Planning	.203	2.50	6	59	.032
Discipline	.125	1.40	6	59	.229
Counterbalancing	.094	1.02	6	59	.421
Topic familiarity	.234	3.00	6	59	.013

As displayed in Table 5.1, the covariate, C-test (proficiency), has significant effects on general task performance (p = .000). The rest of the table clearly shows that, when proficiency is controlled for, there are no differences between male and female students (p = .984), or between computer majors and medicine majors (p = .229), or between those who did the familiar tasks first and those who did the unfamiliar task first (p = .421). This result is encouraging in that it indicates that the extraneous variables were well controlled, and in particular, the problem of any potential performance differences between the computer and medicine majors can be attributed simply to their proficiency difference, but not their discipline *per se* (p = .229). Therefore, we don't need to look at the two discipline groups separately and the original study design can be kept intact. At the same time, the other two target independent variables (topic familiarity, p = .013 and planning, p = .032) reached significance, which means their effects are independent of proficiency (because in this MANCOVA, proficiency has been keep constant). These results lend support to the following sections to proceed with its original goal – the investigation of the effects that topic familiarity, planning and proficiency levels exert on task performance.

5.2 MANOVA results: A general picture of task performances

The present study involves one within-subject independent variable (topic familiarity) and two between-subject independent variables (planning and proficiency), making it a split-plot factorial design (cf., Gardner, 2001, p.127 - 153). Moreover, five categories of dependent variables, namely *fluency, accuracy, complexity, lexis,* and *formality,* are adopted as indices to gauge task performance, all of which, except the lexis, are broken down into various aspects of the same construct (see Table 3.7 above for details). As such, the best statistical procedure that can take into account multiple dependent variables in a split-plot factorial design is a repeated measures multivariate analysis of variance (MANOVA).

Two immediate advantages of a repeated measures MANOVA fit this study well. Firstly, it deals with a research design including both within-subject and between-subject variables. Secondly, it takes care of all dependent variables simultaneously to produce a single multivariate effect statistic, which tells us whether on the whole a certain independent variable does have a significant effect for all dependent variables. According to Meyers, Gamst, and Guarino (2006), all dependent variables in a multivariate analysis should be moderately correlated with a correlation coefficient r between 0.2 and 0.4 being the best. The logic for such a requirement is that if the variables are not correlated, it would be pointless to lump them together and consider them as a whole. Meanwhile, if the variables are highly correlated, they could be regarded as measuring the same construct, and statistically might also fall prey to a multicollinearity situation. An emerging problem here is that many of the variables under the same category are quite highly correlated. Take speech rate (words per minute), mean length of run and phonation time as examples.

Table 5.2 Pearson correlation coefficients between speech rate, MLR and phonation time in the familiar topics (N=80).

	1.MLR	2. WPM	3. Phon.Time
1. Mean Length of Run	1	.842**	.618**
2. Words per Minute		1	.724**
3. Phonation Time			1

** $p < .01$

These three measures variables are supposed to measure different aspects of the same construct: *fluency*, but they are highly correlated (r = .618 - .842), as shown in table 5.2 above. Therefore putting these three variables into the same multivariate analysis would be deemed inappropriate.

To solve the problem, only one or a very small number of dependent variables from each general category, where they all moderately correlate with one another, are chosen as representative for the multivariate analysis. Researchers (e.g., Freed, 2000; Skehan 2001; Tavakoli and Skehan 2005) distinguished the temporal aspects of fluency, under the umbrella term of "breakdown fluency", from the re-organizational aspects of fluency, i.e., the "repair fluency". Therefore, speech rate (words per minute) was chosen to represent the former and total repairs, the latter. In addition to the classic subordination complexity measure, "words per AS unit" was included in the analysis for its representativeness of clausal length, as it is assumed to be a different aspects of syntactic complexity (Ortega, Iwashita, Norris, and Rabie, in preparation). The three lexical measures are inherently different measures which may not reflect a same construct. Therefore they are all drawn into the MANOVA procedure. The F-score is also included to represent the construct of "formality". The selected representative variables are shown in Table 5.3.

The results of the repeated measures MANOVA are shown in Table 5.4. While the significance level (p value) allows us to know the likelihood that the experimental results differ from chance expectations, we are still lacking in

Table 5.3 Variables chosen for the multivariate analysis.

Variable Categories	Variable names
Fluency	1. Words per minute (Breakdown fluency)
	2. Total repairs per 100 words (Repair dysfluency)
Accuracy	Ratio of error-free clauses
Complexity	1. Ratio of clauses per AS unit (Subordination)
	2. Words per AS unit (Length of AS units)
Lexis	1. *Lambda* (lexical sophistication)
	2. *D* (lexical diversity)
	3. ratio of content words to total words (lexical density)
Formality	F-score

Table 5.4 Results of the repeated measures MANOVA (multivariate tests)

Effects	Pillai's Value	F	BGdf	WGdf	Sig.	Cohen's *d*
Between-subjects effects						
Planning	.53	8.60	9	68	.000	.47
Proficiency	.34	3.92	9	68	.000	.32
Planning×Proficiency	.19	1.75	9	68	.093	/
Within-subjects effects						
Familiarity	.53	8.65	9	68	.000	.47
Familiarity×planning	.26	2.71	9	68	.009	.26
Familiarity×proficiency	.11	.93	9	68	.502	/
Familiarity×planning×Proficiency	.12	.99	9	68	.457	/

Note: BGdf = between-group degree of freedom, WGdf = within-group degree of freedom.

information about the magnitude of any significant effect. Cohen's d is a value commonly used to show the size of an experimental effect (Thalheimer and Cook, 2002). Put simply, the p value tells us whether there is an effect, and the Cohen's d value tells us how big the effect is. All results presented in this thesis will include not only the means for each cell, but also the significance level (p[5]) and the effect size (Cohen' d[6]), with Cohen's d calculated with the formula in Thalheimer and Cook (2002).

As shown in Table 5.4 all three independent variables have very significant overall effects on the five categories of dependent variables ($p < .001$ in all three), which licenses us to proceed to the univariate test for each specific dependent variable. Specifically familiarity and planning achieve medium effect sizes ($d = .47$ for both), while proficiency appears to be less powerful as evidenced by the small effect size ($d = .32$). On the whole, familiarity has a significant interaction with planning ($p = .009$), with a quite small effect size though ($d = .26$). Univariate test results (a repeated measures ANOVA) for each dependent variable will be presented next in the order of *total words*, *breakdown fluency*, *repair fluency accuracy*, *complexity*, *lexis*, and *formality*, with a range of different variables subsumed under each of these major categories. In each section, first a few tables will sum up the basic statistics of each cell, followed by some explanations of the

---

[5] The conventional $p<0.05$ level is set as the cut-off line for a significant effect.

[6] According to Cohen (1992), the effect size of .20 is small, .50 is medium, and .80 is large. Halheimer and Cook (2002) further divide effect sizes into 6 levels: negligible effect: >=-0.15 and <.15; small effect:>=.15 and <.40; medium effect: >=.40 and <.75; large effect: >=.75 and <1.10; very large effect: >=1.10 and <1.45; huge effect: >1.45.

results. Where there is any significant interaction between the independent variables, a graph will also be presented to illustrate the findings visually.

5.3 Number of Total Words

Two types of "total words" will be reported here. "Raw words" is a measure of all utterances a speaker produces, and 'pruned words' includes all utterances except filled pauses (e.g., "er", "mmm", and "ah"), pseudo filled pauses (e.g., "well", "like", and "actually" in some circumstances), false starts, reformulations, replacements, and repetitions, thus becoming a more genuine measure of total words produced.

Table 5.5 Means of total words in a task (N=40 in each cell)

	Familiarity		Planning		Proficiency	
	Fam.	Unfam.	Planned	Non-planned	High	Intermed.
Raw words	360.36 (166.21)	284.05 (117.64)	348.46 (154.85)	295.95 (123.87)	326.87 (135.27)	317.44 (149.35)
Pruned Words	300.84 (138.47)	229.61 (94.11)	297.85 (126.02)	233.60 (96.89)	267.95 (112.46)	262.50 (120.89)

Note: Standard deviations in ().

A repeated measures ANOVA was conducted to test the effects of topic familiarity, planning and proficiency on the two types of total words (The same will apply to every dependent variable hereafter). The results (Table 5.5) show that participants produced longer accounts on the more familiar topics, $F(1, 76) = 28.33$, $p = .000$, Coheh's $D = .53$ for raw words, and $F(1, 76) = 35.03$, $p = .000$, Cohen's $d = .61$ for pruned words. The opportunity for planning time seems to be a less powerful means in pushing learners to say more, as a significant effect is reached only with total pruned words ($F(1, 76) = 7.60$, $p = .007$, Cohen's $d = .44$), which is an indication that participants produce fewer "useless" utterances after planning. A comparison of effect sizes further supports the argument that familiarity with a certain topic pushes participants to use more words than does planning. Proficiency does not have any effect on the number of words. Neither are there any interactions. Therefore, the effects of topic familiarity and planning here are quite straightforward.

Summary of total words results: *Topic familiarity* is a favorable condition for saying more words on a task. *Planning* time only has an effect for pruned speech. *Proficiency* does not show any effect in this regard.

5.4 Breakdown Fluency

The main effects for each independent variable and their interactions (if any) from section 5.4 to 5.9 will be reported in separate tables, though the results were in fact conducted in one repeated measures ANOVA for each dependent variable. In this part, the first three tables report on the effects of the three independent

Table 5.6 The effects of topic familiarity on breakdown fluency

	Topic Familiarity						Cohen's *d*
	Fam	Unfam	F	BGdf	WGdf	Sig.	
Words per Minute	96.30 (23.02)	90.47 (26.33)	25.78	1	76	.000	.26
Phonation Time	.80 (.09)	.77 (.10)	24.00	1	76	.000	.35
Mean length of run	5.26 (1.57)	4.99 (1.76)	6.05	1	76	.016	.17
Filled pause number	9.71 (4.77)	10.55 (6.00)	3.03	1	76	.085	/
Mid-clause pause number	9.73 (6.20)	12.13 (7.33)	27.50	1	76	.000	.38
End-of-clause pause number	6.53 (2.09)	6.82 (2.52)	2.47	1	76	ns	/
Mid-clause silence total	8.51 (7.86)	12.35 (13.41)	14.80	1	76	.000	.38
End-of-clause silence total	5.73 (2.91)	6.41 (4.04)	6.33	1	75	.014	.19
Mid-clause pause length	.79 (.31)	.87 (.35)	5.73	1	76	.019	.28
End-of-clause pause length	1.71 (.61)	1.78 (.81)	.92	1	76	ns	/

Notes: 1. Standard deviation in (). 2. All pause number and silence measures, including filled pauses, are standardized by calculating their occurrence per 100 words. 3. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

variables for the 10 breakdown fluency measures, followed by a fourth tables which shows their interactions.

Table 5.6 above could be summed up in two points. First, it seems that topic familiarity displays an overall effect on most (8 out of 10) breakdown fluency measures. Being familiar with a certain knowledge domain drives the participants to speak at a higher speech rate (p = .000, d = .26), with more time spent on speaking and thus less silence (phonation time, p = .000, d = .35), in a longer stretch of words before encountering any pauses, repairs or filler (mean length of run, p = .016, d = .17). Familiarity with a topic also helps to reduce the number (p = .000, d = .38) as well as the average length (p = .019, d = .25) of pauses, and the total silence (p = .000, d = .38), in the middle of a clause. In addition, topic familiarity is able to shorten the total silence time between two clauses (end-of-clause silence, p = .014, d = .19). However, filled pauses, and the number and length of pauses at the end of clauses seem unaffected by topic familiarity (p > .05 in both cases). The second feature this table reveals is the consistently small effect sizes in all measures contrasted with the wider range of significance values. None of the effect sizes reaches the medium level, which is a signal that while topic familiarity leads to higher fluency, it does not work very powerfully.

The effects of planning, as shown in Table 5.7 below, are quite similar to those of topic familiarity, except that planning achieves a significant impact on more measures (9 out 10) with a larger magnitude of the effects (generally bigger effect sizes). The opportunity to plan prior to speaking raises the speech rate (p

Table 5.7 The effects of planning on breakdown fluency

	Planning						Cohen 's d
	Planned	Unplanned	F	BG df	WG df	Sig.	
Words per Minute	102.52 (22.45)	84.24 (23.51)	13.35	1	76	.000	.58
Phonation Time	.82 (.07)	.75 (.10)	14.34	1	76	.000	.62
Mean length of run	5.48 (1.85)	4.77 (1.38)	4.10	1	76	.046	.32
Filled pause number	9.90 (5.49)	10.36 (5.12)	.17	1	76	ns	/
Mid-clause pause number	8.61 (5.27)	13.24 (7.31)	11.66	1	76	.001	.54
End-of-clause pause number	6.32 (2.22)	7.05 (2.34)	2.40	1	76	ns	/
Mid-clause silence Total	6.46 (4.83)	14.39 (13.09)	14.73	1	76	.000	.61
End-of-clause silence total	5.10 (2.80)	7.08 (3.83)	7.71	1	75	.007	.59
Mid-clause pause length	.70 (.33)	.96 (.40)	19.74	1	76	.000	.71
End-of-clause pause length	1.53 (.42)	1.96 (.86)	9.02	1	76	.004	.64

Notes: 1. Standard deviation in (). 2. All pause number and silence measures, including filled pauses, are standardized by calculating their occurrence per 100 words. 3. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.8 The effects of proficiency on breakdown fluency

	Proficiency		F	BG df	WG df	Sig.	Cohen's d
	High	Intermed					
Words per Minute	96.79 (25.26)	89.97 (23.90)	1.86	1	76	ns	/
Phonation Time	.80 (.08)	.77 (.11)	1.92	1	76	ns	/
Mean length of run	5.28 (1.51)	4.98 (1.80)	.73	1	76	ns	/
Filled pause number	10.08 (5.85)	10.18 (4.96)	.007	1	76	ns	/
Mid-clause pause number	10.17 (5.96)	11.69 (7.47)	1.26	1	76	ns	/
End-of-clause pause number	6.15 (1.97)	7.22 (2.48)	5.08	1	76	.027	.48
Mid-clause silence total	9.35 (9.25)	11.51 (11.77)	1.10	1	76	ns	/
End-of-clause silence total	5.63 (3.12)	6.54 (3.79)	1.64	1	75	ns	/
Mid-clause pause length	.80 (.29)	.86 (.36)	1.15	1	76	ns	/
End-of-clause pause length	1.77 (.68)	1.72 (.74)	.13	1	76	ns	/

Notes: 1. Standard deviation in (). 2. All pause number and silence measures, including filled pauses, are standardized by calculating their occurrence per 100 words. 3. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

$= .000$, $d = .58$), phonation time ($p = .000$, $d = .62$), as well as the mean length of run ($p = .046$, $d = .32$). Planning reduces the number of mid-clause pauses ($p = .001$, $d = .54$), though not the number of end-of-clause pauses ($p > .05$). The reduction also occurs to the amount of silence ($p = .000$, $d = .61$) and the average pause length ($p = .000$, $d = .71$) in the middle of a clause, and the amount of silence ($p = .007$, $d = .59$) as well as the average pause length ($p = .004$, $d = .64$) at the end-of-clause positions. Like familiarity, planning seems not to have significant influence on the filled pauses ($p > .05$).

Rather counter-intuitively, proficiency appears to be irrelevant to all but one measure. However, the one and only significant influence reached by proficiency is a very interesting occurrence in that the number of pauses at clause boundaries is one of few measures that neither topic familiarity nor planning can exert their influence on, whereas proficiency happens to fill this vacancy ($p = .027$), with a medium Cohen's *d* value ($d = .48$) indicating a considerable effect. Moreover, we might conclude at this point that filled pauses, e.g., "er", "mmm", and "eh", appear to be a more inherent feature in L2 speaking as it is resistant to the influence of familiarity, planning, or proficiency.

Table 5.9 below shows seven significant interactions that by default could be categorized into three types. The first type, i.e., the familiarity by planning interaction, echoes the MANOVA test results (Table 5.4). This type of interaction occurs in five out of the 10 breakdown fluency measures, which might at least partially explain the multivariate results (see Table 5.4, p.98, the only significant

Table 5.9 *p* values of significant interactions among topic familiarity, planning and proficiency

	Interactions			
	Fam×Plan	Fam×Prof	Plan×Prof	Fam×Plan×Prof
Words per Minute	.001	ns	ns	.034
Phonation Time	.000	ns	ns	ns
Mean length of run	ns	ns	ns	ns
Filled pause number	ns	ns	ns	ns
Mid-clause pause number	.005	ns	ns	ns
End-of-clause pause number	ns	.017	ns	ns
Mid-clause silence Total	.004	ns	ns	ns
End-of-clause silence total	.026	ns	ns	ns
Mid-clause pause length	ns	ns	ns	ns
End-of-clause pause length	ns	ns	ns	ns

Note: All pause number and silence measures, including filled pause number, are standardized by calculating their occurrence per 100 words.

interaction effect in the MANOVA results is the Familiarity × Planning effect, $p = .009$).

The following Table 5.10 shows the means of the five dependent variables for the topic familiarity × planning interaction. A general trend is clearly displayed in this table – planning greatly compensates for the unfavorable condition induced by the unfamiliar topics. The significant difference in breakdown fluency between familiar and unfamiliar topics is reduced to almost non-existence when planning is allowed, especially in speech rate (p = .001), phonation time (p = .000) and the end-of-clause silence total (p = .026). Planners were always able to speak faster, pause less frequently in the middle of a clause, and reduce the silence time at both mid-clause and end-of-clause positions, regardless of their familiarity with the topics. Following from this interaction, two points are worth mentioning.

Table 5.10 Means of words per minute, phonation time, mid-clause pause number, mid-clause silence total and end-of-clause silence total

	Topic familiarity	
	Familiar	Unfamiliar
Words per minute		
Planned	103.47 (20.88)	101.58 (24.01)
Unplanned	89.12 (23.06)	79.35 (23.96)
Phonation time		
Planned	.83 (.08)	.82 (.07)
Unplanned	.78 (.09)	.72 (.11)
Mid-clause pause number		
Planned	8.08 (5.16)	9.15 (5.38)
Unplanned	11.37 (6.76)	15.11 (7.86)
Mid-clause silence total		
Planned	6.04 (4.86)	6.89 (4.81)
Unplanned	10.97 (9.44)	17.81 (16.73)
End-of-clause silence total		
Planned	5.06 (2.40)	5.13 (3.19)
Unplanned	6.48 (3.24)	7.72 (4.42)

Notes: 1. Standard deviation in (). 2. Dependent variables in italics. 3. N=40 in each cell.

Firstly, if we compare, in the above Table 5.10, the difference between the familiar and the unfamiliar tasks in the planned condition, and then between the planned and the unplanned speeches in the familiar tasks, we may reach an agreement that planning is capable of narrowing the gap between two familiarity types to a much greater extent than familiarity can do between planned and unplanned tasks. Secondly, planning helps to improve fluency in both familiar and unfamiliar tasks, but obviously the unfamiliar tasks benefit much more. These results emerge more clearly from the following graphs (Figure 5.1 to 5.5).

Figure 5.1 Familiarity×planning interaction in speech rate

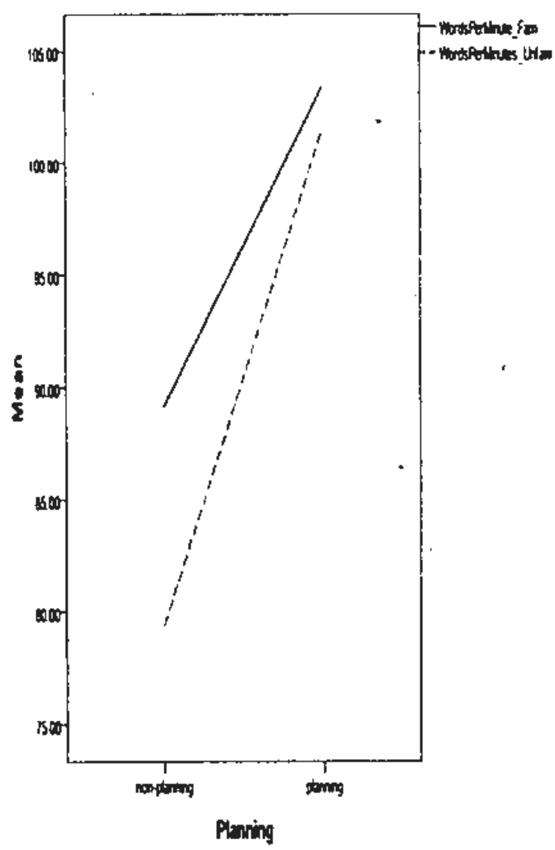


Figure 5.2 Familiarity×planning interaction in phonation time

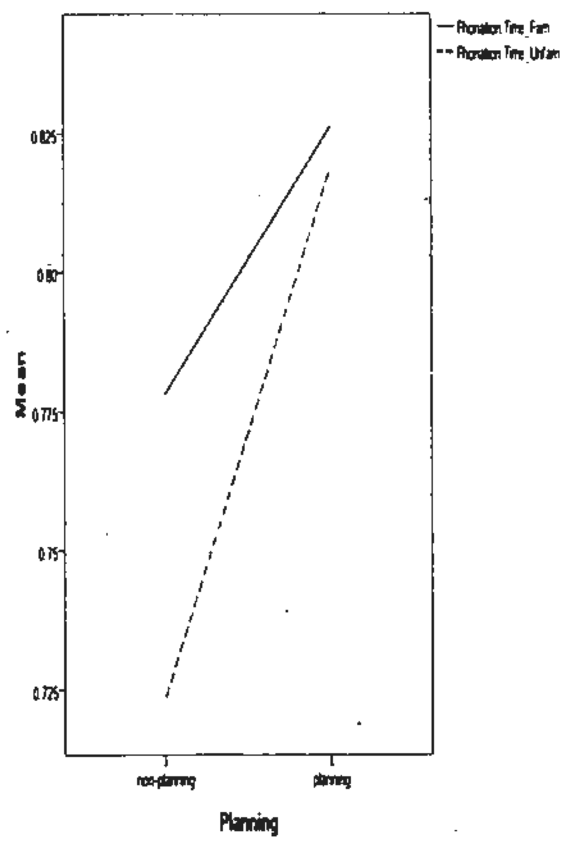


Figure 5.3 Familiarity×planning in number of mid-clause pauses (per 100 words)

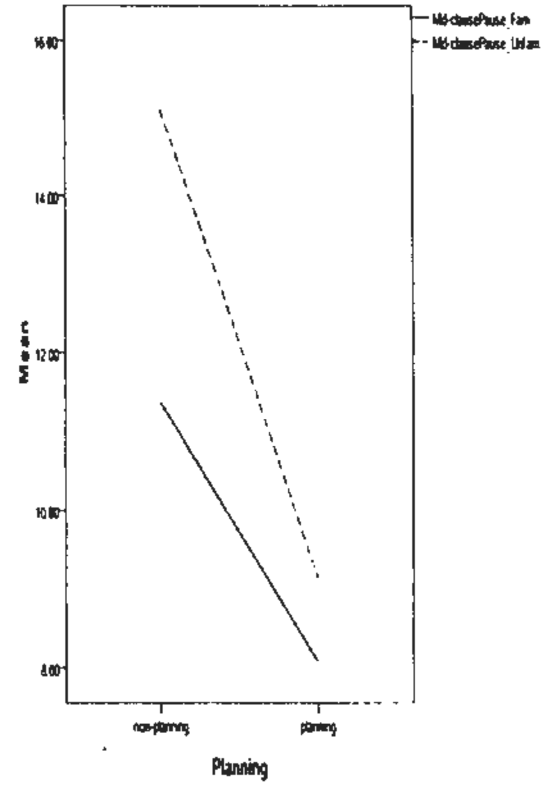


Figure 5.4 Familiarity×planning interaction in mid-clause silence total (per 100 words)

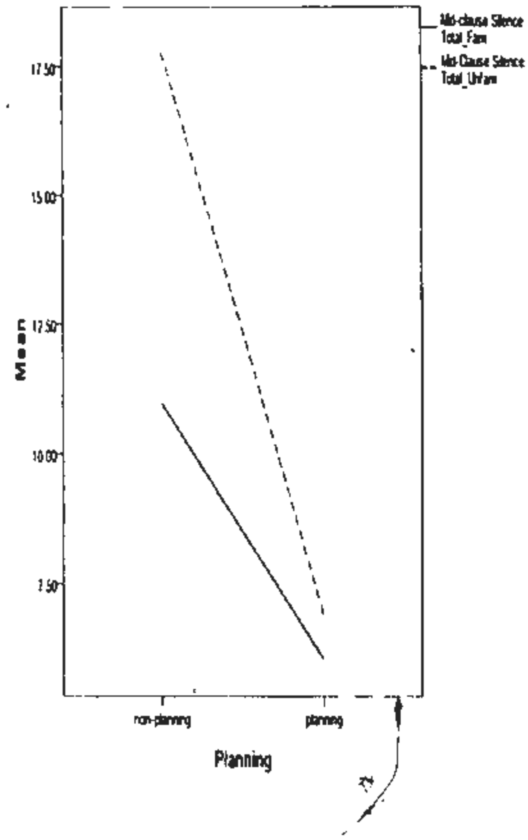
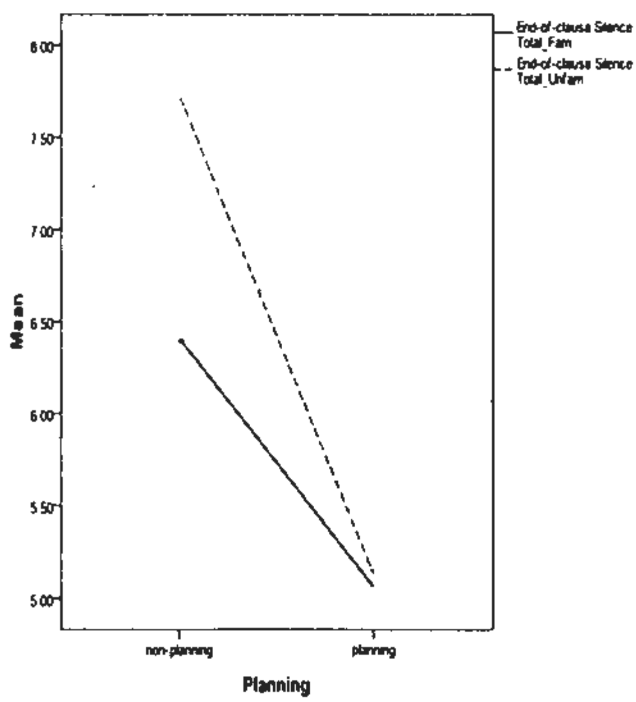


Figure 5.5 Familiarity×planning interaction in end-of-clause silence total (per 100 words)



The second type of interaction occurs between topic familiarity and proficiency in terms of the number of end-of-clause pauses (p = .017). Two notable points stand out from Table 5.11 below. On the one hand, participants with a higher proficiency (6.24 and 6.07 times per 100 words for the familiar and unfamiliar topics respectively) always paused at the end of a clause less than the intermediate ones (6.88 and 7.58 times) did, whether on familiar or unfamiliar topics. On the other hand, lower proficiency learners appeared to be more vulnerable to the influence of familiarity as they paused more in the unfamiliar tasks where their higher proficiency counterparts paused slightly less, hence the interaction, as illustrated in Figure 5.6 below. If we compare end-of-clause pauses to mid-clause pauses (see Table 5.12 after Figure 6), there emerges an interesting trend, arithmetically at least, that proficiency levels affect where to pause. In handling the unfamiliar topics, high proficiency participants produced more pauses in the middle of, but not at the end of, a clause, whereas intermediate participants had to pause at both positions. This suggests that mid-clause pausing is an intrinsic phenomenon in L2 speaking irrespective of proficiency levels but end-of-clause pausing is more sensitive to the influence of proficiency.

The third type of interaction is a complex three-way interaction involving all three independent variables for speech rate (p = .034). There is a two-way (Familiarity × Planning) interaction, as illustrated in Table 5.10 and Figures 5.1 - 5.5 above, which indicates that planning can greatly mitigate the difference

Table 5.11 Means of end-of-clause pause number (per 100 words)

	Topic familiarity	
	Familiar	Unfamiliar
Proficiency		
Intermediate	6.88 (2.09)	7.58 (2.86)
High	6.24 (2.06)	6.07 (1.88)

Notes: 1. Standard deviations in (). 2. N= 40 in each cell.

Figure 5.6 Familiarity × proficiency interaction in end-of-clause pause number (per 100 words)

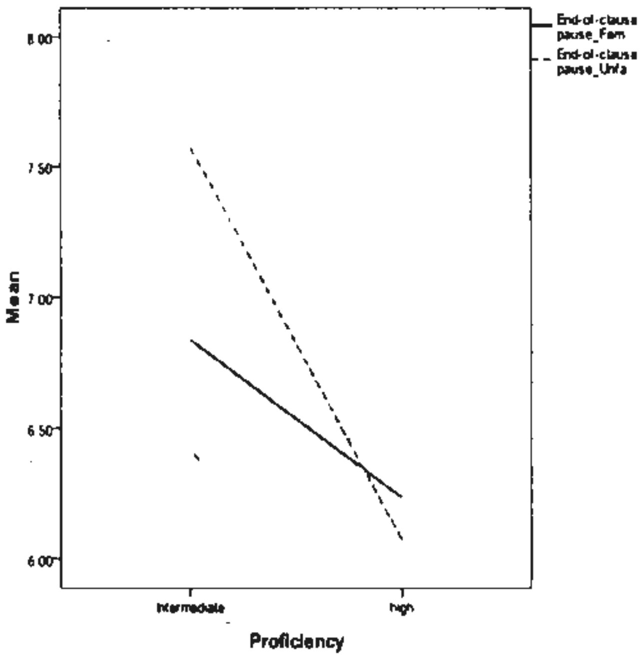


Table 5.12 Means of mid-clause pause number (per 100 words)

	Topic familiarity	
	Familiar	Unfamiliar
Proficiency		
Intermediate	10.27 (7.15)	13.10 (7.79)
High	9.17 (5.11)	11.15 (6.10)

Notes: 1. Standard deviations in (). 2. N= 40 in each cell.

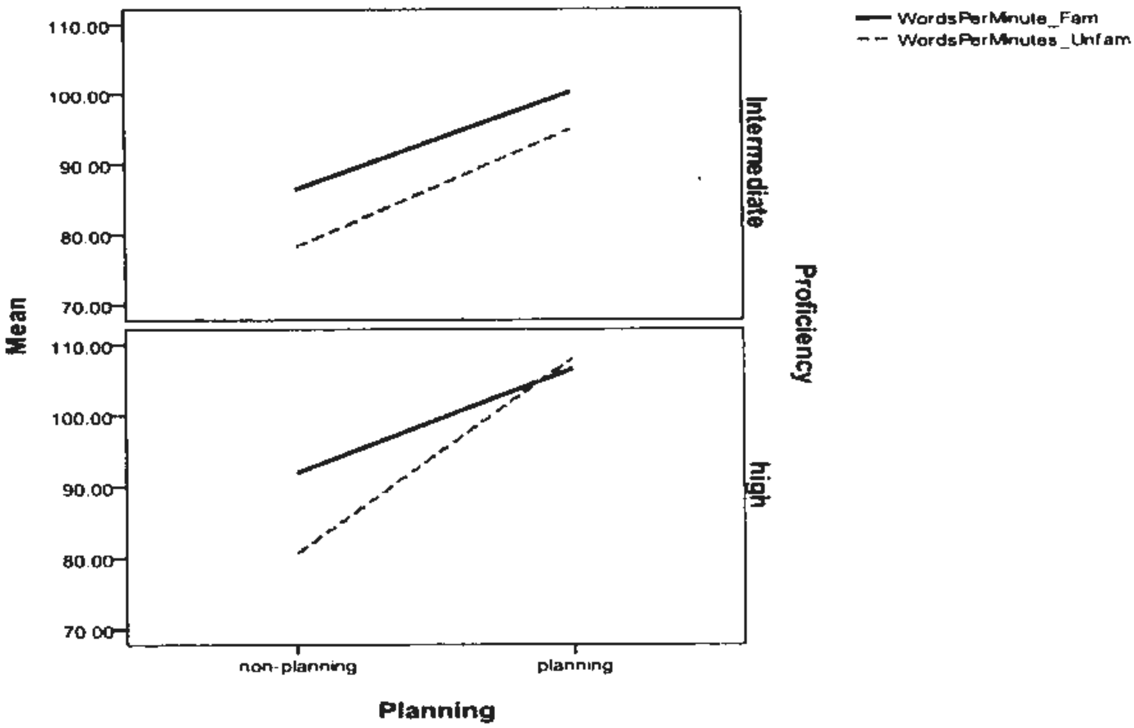
between the familiar and the unfamiliar topic regarding speech rate. However, this three-way interaction reminds us that such an interactive effect is modified by proficiency level. As shown in the Table 5.13 below, a clear interaction between familiarity and planning exists in the high proficiency participants where they spoke even slightly faster on the unfamiliar topic (108.10 words per minute) than the familiar one (106.68) when given planning time. However, such an effect cannot be found in the intermediate participants as they always speak faster on the more favorable conditions, namely when planning time or familiar topics were provided. As illustrated by the two almost parallel lines in Figure 5.7 (upper part) below, planning time does not narrow the gap between the two familiarity conditions but improves them to the same extent.

Table 5.13 Means of words per minute in three-way interaction

		Topic familiarity	
		Familiar	Unfamiliar
Proficiency	Planning		
Intermediate	Unplanned	86.35 (21.83)	78.13 (23.90)
	Planned	100.35 (21.00)	95.06 (24.76)
High	Unplanned	91.90 (24.48)	80.58 (24.59)
	Planned	106.68 (20.81)	108.10 (21.93)

Note: 1. Standard deviations in (). 2. N=20 in each cell.

Figure 5.7 Familiarity×planning×proficiency Interaction in speech rate



Summary of breakdown fluency results:

1) Main effects: *planning* shows very wide effects over almost all breakdown fluency measures, with medium to large effect sizes, proving itself as a useful means in improving the temporal-acoustic aspects of speaking. *Topic familiarity* largely resembles planning in most measures, but it appears less powerful as evidenced by not only the fewer measures it has effects on, but also the much smaller effect sizes it displays. *Proficiency*, somewhat surprisingly, seems ineffective for promoting better fluency, and probably is very much overridden by the other two independent variables.

2) Interactions: Firstly, five interaction effects between *planning* and *familiarity* can be observed, all pointing to a pattern that the dysfluency invoked

by unfamiliarity with a certain topic could be compensated for by pre-task planning. In five out of the ten breakdown fluency measures, planners achieved an almost equally fluent delivery of speech in spite of their familiarity levels with the topics whereas non-planners spoke significantly better on the more familiar topic. Secondly, the *familiarity* × *proficiency* interaction in the number of end-of-clause pause indicates that intermediate participants had to pause more frequently at the end of a clause when speaking on an unfamiliar topic than a familiar topic, whereas the higher proficiency participants seemed almost unaffected in this respect. The three-way *planning* × *familiarity* × *proficiency* interaction found in 'words per minute' shows that the more advanced students made better use of planning time to improve their speech rate on the unfamiliar topic than the intermediate students whose speech rate on the unfamiliar topic was still significantly lower than the familiar one even if given planning time.

5.5 Repair Fluency

In this section, four repair fluency measures will be reported as individual variables. An additional variable, 'total repairs' which is the sum of the four measures, will also be included here as a potentially more all-encompassing index for repair fluency.

Table 5.14 shows that topic familiarity only helps to reduce the number of

Table 5.14 The effects of topic familiarity on repair fluency

	Topic Familiarity			BG	WG		Cohen'
	Fam.	Unfam.	F	df	df	Sig.	s *d*
False starts	1.38	1.63	2.40	1	76	ns	/
	(1.28)	(1.35)					
Reformulations	1.39	1.62	2.99	1	76	.088	/
	(1.00)	(1.25)					
Replacements	.95	1.15	3.21	1	76	.077	/
	(.79)	(.97)					
Repetitions	3.94	4.72	12.19	1	76	.001	.31
	(2.69)	(3.36)					
Total repairs	7.55	9.11	19.22	1	76	.000	.40
	(3.63)	(4.92)					

Note: 1. All measures were standardized by calculating their occurrence per 100 words. 2. Standard Deviations in (). 3. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

repetitions (p = .001), with a small effect size (d = .31). Though not reaching significance, the means of other three variables are on the predicted direction, which explains why topic familiarity significantly works for fewer total repairs (p = .000) at a medium effect size magnitude (d = .40).

Compared to the effects of topic familiarity, ten-minute planning (Table 5.15) is able to not only lower the total number of repairs (p = .000, d = .75), but also to display significant effects on all four single repair measures that constitute the total, though the pattern is not that consistent. Planning helps participants with fewer false starts (p = .000, d = 1.02), reformulations (p = .001, d = .53) and repetitions (p = .000, d = .60), but it also drives them to use more replacements (p = .008, d =

Table 5.15 The effects of planning on repair fluency

	Planning		F	BG df	WG df	Sig.	Cohen's *d*
	Unplan'd	Planned					
False starts	2.15 (1.40)	.85 (.80)	41.32	1	76	.000	1.02
Reformulations	1.84 (1.26)	1.16 (.84)	11.11	1	76	.001	.53
Replacements	.84 (.81)	1.26 (.90)	7.36	1	76	.008	.43
Repetitions	5.40 (3.44)	3.16 (2.03)	14.31	1	76	.000	.60
Total repairs	10.23 (4.32)	6.44 (3.29)	22.04	1	76	.000	.75

Table 5.16 The effects of proficiency on repair fluency

	Proficiency		F	BG df	WG df	Sig	Cohen's *d*
	High	Intermediate					
False starts	1.50 (1.29)	1.50 (1.35)	.00	1	76	ns	/
Reformulations	1.60 (1.28)	1.40 (.95)	1.00	1	76	ns	/
Replacements	1.08 (.85)	1.02 (.92)	.14	1	76	ns	/
Repetitions	4.29 (3.32)	4.27 (2.74)	.00	1	76	ns	/
Total repairs	8.47 (4.57)	8.20 (4.00)	.23	1	76	ns	/

Note for table 5.15 and 5.16: 1. All measures were standardized by calculating their occurrence per 100 words. 2. Standard Deviations in (). 3. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

.43). Similar to those with breakdown fluency, the effect sizes produced by planning for repair fluency range from medium to large, much bigger than those for familiarity.

A caveat has to be made concerning the nature of repair fluency (or rather, dysfluency). Repairs appear to be, at first sight, an interruption to the flow of smooth speech that has certain resemblance to breakdowns in this regard. However, what differentiates repairs from breakdowns is that repairs are indicative of the effort on the part of the learners to fill the silence and to perhaps look for better utterance, which should not be viewed negatively. In the six "Ealing studies" conducted by Skehan and Foster (C.f., Skehan & Foster, 2007), the opportunity to plan before tasks was able to reduce pauses (few breakdowns) but it nevertheless pushed learners to produce more repairs. In comparison, this study, though showing that other repairs were less frequent after planning, the number of replacement significantly increased. This is interesting in that repairs seem to behave a bit differently from other fluency measures and has in itself a potential of positive effects for the learners to do re-organization and revision of their language.

Also resembling breakdown fluency, proficiency seems to exert no effect on repair fluency. What stands in contrast with breakdown fluency are the interaction effects. Since no interaction could be indentified in repair fluency, the main effects of topic familiarity and planning as two favorable conditions in general for better repair fluency are confirmed.

Summary of repair fluency results:

Topic familiarity reduces repetitions as well as the total number of repairs. *Planning* reduces false starts, reformulations, repetitions and the total number of repairs, but increases replacements. Compared with familiarity, planning shows bigger effect sizes as well as influences on a wider range of repair measures. *Proficiency* does not have any significant effects on repair fluency variables. There is no interaction found between any independent variables.

5.6 Accuracy

This section reports on the three accuracy measures. First, three tables sum up the main effects of topic familiarity, planning and proficiency respectively, with accompanying description. Then, one interaction effect between familiarity and planning will be presented.

Table 5.17 The effects of topic familiarity on accuracy

	Topic familiarity			BG	WG		Cohe
	Familiar	Unfam.	F	df	df	Sig.	n's *d*
Error-free clause ratio	.544 (.13)	.517 (.14)	5.60	1	76	.020	.22
70% Accuracy Clause length	3.73 (2.27)	3.68 (2.17)	.04	1	76	ns	/
Errors per 100 words	6.86 (2.46)	7.71 (2.61)	17.81	1	76	.000	.38

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.18 The effects of planning on accuracy

	Planning		F	BG df	WG df	Sig.	Cohen 's *d*
	Unplan'd	Planned					
Error-free clause ratio	.524 (.13)	.537 (.14)	.25	1	76	ns	/
70% Accuracy Clause length	3.61 (1.96)	3.79 (2.63)	.20	1	76	ns	/
Errors per 100 words	7.64 (2.45)	6.92 (2.59)	2.45	1	76	ns	/

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.19 The effects of proficiency on accuracy

	Proficiency		F	BG df	WG df	Sig.	Cohen's *d*
	High	Intermediate					
Error-free clause ratio	.586 (.13)	.475 (.11)	19.92	1	76	.000	.69
70% Accuracy Clause length	4.4 (2.49)	3.0 (1.66)	12.82	1	76	.001	.57
Errors per 100 words	6.16 (2.25)	8.41 (2.30)	23.54	1	76	.000	.77

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

As shown in tables 5.17, 5.18, and 5.19 below, topic familiarity appears to push participants to achieve a higher ratio of error-free clauses (p = .020, d = .22), and reasonably fewer errors per 100 words (p = .000, d = .38), with small effect sizes though. Being familiar with a topic, however, is not able to help learners to produce longer clauses where at least 70% of these clauses are correct (70% accuracy clause length, p > .05). Planning does not show an effect on any of the measures here (p > .05 in all three measures). Proficiency is a strong driving force for accuracy as evidenced by the medium to large effect sizes. The more advanced participants performed with longer 70% accuracy clauses (p = .001, d = .57), in addition to their ability of having higher error-free ratio (p = .000, d = .69) as well as a smaller number of total errors per 100 words (p = .000, d = .77), when compared with their intermediate counterparts.

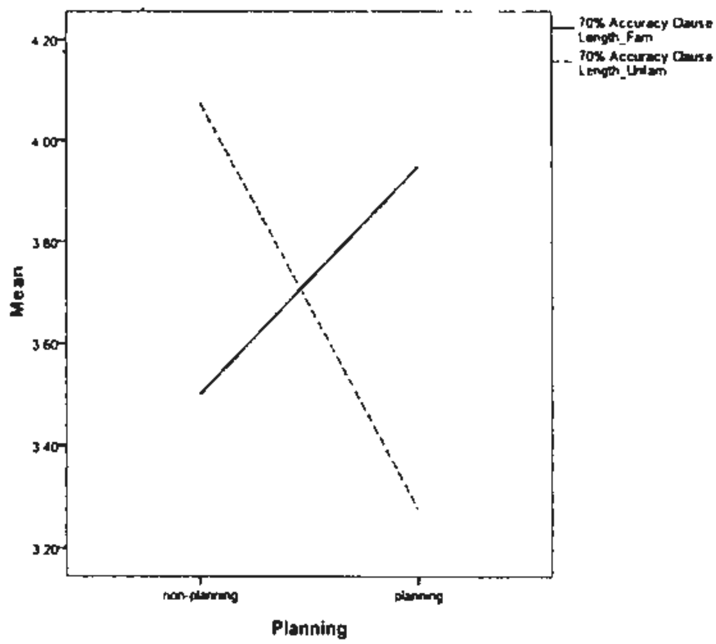
Table 5.20 Means of 70% accuracy clause length

	Topic familiarity	
	Familiar	Unfamiliar
Planning		
non-planning	3.50 (1.80)	4.08 (2.13)
planning	3.90 (2.67)	3.28 (2.16)

Note: Standard deviations in ().

There is an interaction between familiarity and planning in the measure of '70% accuracy clause length', which is a measure considering both error numbers and clause length. A closer examination of the means in Table 5.20 above shows that the non-planners were more accurate whenperforming the unfamiliar topic than when doing the familiar topic, whereas the planners were exactly the opposite, as is more clearly illustrated in Figure 8. This interaction seems difficult to explain in that it contradicts the other two measures among the non-planners where the unfamiliar topics induce higher accuracy rate. Also hard to disentangle is the puzzle as to why planners scored lower in the more unfamiliar topic, but higher in the familiar topic, than the non-planners.

Figure 5.8 Familiarity×planning Interaction
in 70% accuracy clause length



<u>Summary of accuracy results:</u>

Proficiency is the most important variable that influenced the accuracy levels of performance because it shows clear main effects on all three accuracy measures with medium to large effect sizes. Being *familiar* with a topic can significantly help reduce the number of errors, with small effect sizes though, but it is not able to raise the length of error-free clauses. *Planning* does not seem to affect accuracy in L2 task performance.

A surprising *familiarity × planning* interaction occurs in the '70% accuracy clause length' measure whereby accuracy is raised with unfamiliar topics, but lowered with familiar topics, among the non-planners, which is opposite to the trend of the planners.

5.7 Complexity

As in the previous sections, the following three tables display the effects of the three independent variables on the three complexity measures, followed by some analysis of the results. In addition, there will be a report on three significant interactions.

Tables 5.21, 5.22, and 5.23 give an account of the three different complexity measures. Topic familiarity seems irrelevant to any of the complexity measures ($p > .05$ for all three measures), but two measures, namely 'clauses per AS unit' ($p = .018$ $d = .39$) and 'words per AS unit' ($p = .000$, $d = .81$), are significantly influenced by planning in which planners outperformed non-planners, with small and large effect sizes respectively.

Table 5.21 The effects of topic familiarity on complexity

	Topic familiarity			BG	WG		Cohen
	Familiar	Unfam.	F	df	df	Sig.	's d
Clauses per AS unit	1.74 (.32)	1.73 (.35)	.11	1	76	ns	/
Words per AS unit	12.93 (2.69)	12.43 (3.36)	2.98	1	76	ns	/
Words per Clause	7.11 (.77)	6.97 (.85)	1.71	1	76	ns	/

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom..

Table 5.22 The effects of planning on complexity

	Planning		F	BG df	WG df	Sig.	Cohen's *d*
	Unplan'd	Planned					
Clauses per AS unit	1.67 (.35)	1.81 (.31)	5.81	1	76	.018	.39
Words per AS unit	11.39 (3.36)	13.96 (2.2)	25.78	1	76	.000	.81
Words per Clause	6.95 (.75)	7.13 (.62)	1.53	1	76	ns	/

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.23 The effects of proficiency on complexity

	Proficiency		F	BG df	WG df	Sig.	Cohen's *d*
	High	Intermed.					
Clauses per AS unit	1.79 (.34)	1.68 (.32)	3.44	1	76	.067	/
Words per AS unit	13.49 (2.70)	11.85 (3.15)	10.57	1	76	.000	.52
Words per Clause	7.15 (.81)	6.92 (.82)	2.58	1	76	ns	/

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Participants of higher proficiency spoke with longer AS units than those of lower proficiency (p = .000, d = .52). Though only approaching significance (p

= .067), the *p* value in 'clauses per AS unit' shows a trend in that the advanced learners are probably able to produce a higher subordination ratio than the intermediate ones. In comparison to the effect size for proficiency, planning appears to be a stronger variable in promoting complexity. A bit unexpectedly, the newly developed measure of 'words per clause' does not seem to be sensitive to the influence of familiarity, planning, or proficiency.

Three interactions occur with the complexity measures, as indicated in Table 5.24. First, there is an interaction between familiarity and planning in terms of 'clauses per AS unit'. Judging from Table 5.25 below, we may see that, though planning generally raises complexity scores, non-planners did better on the familiar topic whereas the planners achieved a higher level on the unfamiliar topic. Therefore, the unfamiliar topc appears to be influenced by planning much more than the familiar topic in terms this complexity measure.

Table 5.24 *p* values of significant interactions among topic familiarity, planning and proficiency in complexity measures

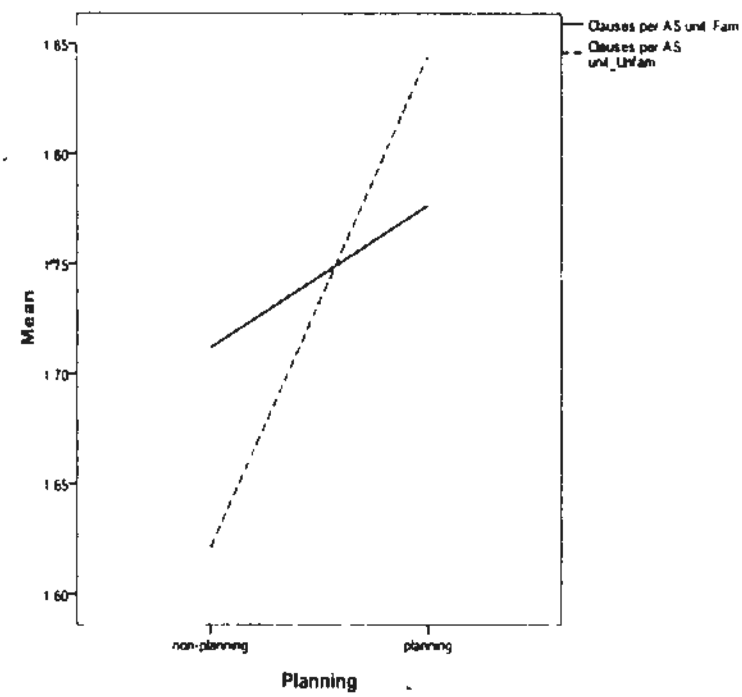
	Interactions			
	Fam×Plan	Fam×Prof	Plan×Prof	Fam×Plan×Prof
Clauses per AS unit	.029	ns	.002	ns
Words per AS unit	ns	ns	.026	ns
Words per Clause	ns	ns	ns	ns

Table 5.25 Means of clauses per AS unit

	Topic familiarity	
	Familiar	Unfamiliar
Planning		
non-planning	1.71 (.36)	1.62 (.33)
planning	1.78 (.28)	1.84 (.33)

Note: 1. Standard deviations in ( ).

Figure 5.9 Familiarity × planning interaction in 'clauses per AS unit'



Also with 'clauses per AS unit', the second interaction is between planning and proficiency (p = .002). Table 5.26 and Figure 5.10 show that, while higher proficiency participants were even slightly lower in complexity after planning (planned 1.77 Vs non-planned 1.82), the intermediate participants clearly produced higher scores when given planning time (planned 1.85 Vs non-planned

1.52).

Table 5.26 Means of clauses per AS unit

	Proficiency	
	High	Intermediate
Planning		
non-planning	1.82 (.40)	1.52 (.20)
planning	1.77 (.27)	1.85 (.34)

Note: 1. Standard deviations in ().

Figure 5.10 Proficiency × planning interaction in 'clauses per AS unit'

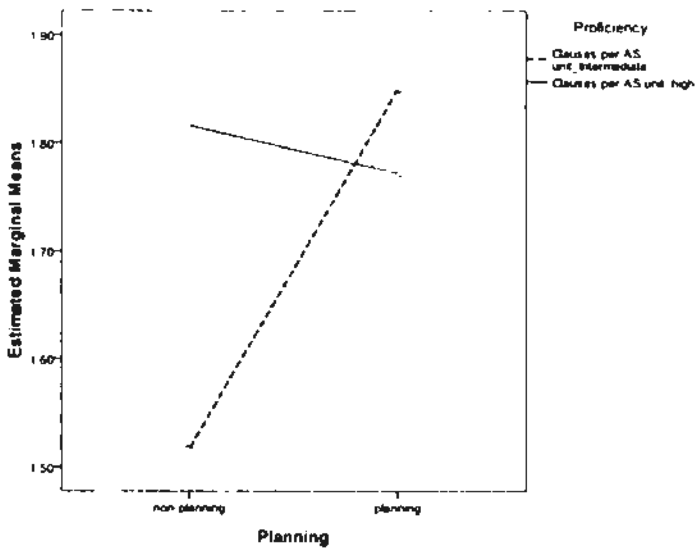
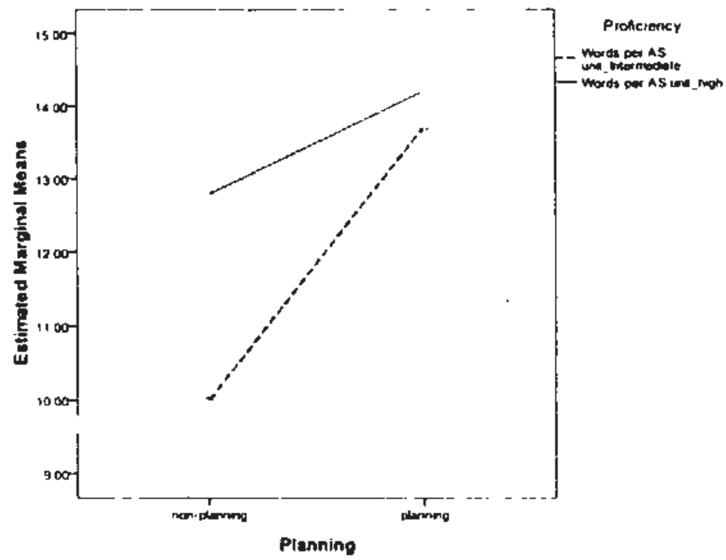


Table 5.27 Means of words per AS unit

	Proficiency	
	High	Intermediate
Planning		
non-planning	12.78 (3.18)	9.99 (2.90)
planning	14.21 (1.93)	13.71 (2.14)

Note: 1. Standard deviations in ().

Figure 5.11 Proficiency × planning interaction in 'words per AS unit'



A third interaction was alos found between proficiency and planning for the measure of 'words per AS unit' (p = .026). Table 5.27 above shows that, though the more advanced participants always scored higher than the intermediate ones, the gap between them is significantly narrowed after planning, as illustrated in figure 5.11 above. Planning, while raising the length of AS units for all, appears to help participants of lower proficiency more than the higher.

<u>Summary of complexity:</u>

1) Main effects: *Topic familiarity* does not influence complexity. The higher proficiency participants produced longer AS units, and almost significantly a higher subordination ratio (clauses per AS unit), than the intermediate ones. *Planning* significantly raises the number of both clauses and words in an AS unit, with large effect sizes than those of proficiency. None of the three independent

variables shows an effect on 'words per clause'.

2) Interactions: Three interactions occur with the complexity measures, out of which a pattern can be extracted that planning is a strong mediating variable that is able to mitigate the difference between the familiar and unfamiliar topics, and between the higher and lower proficiency levels, in terms of complexity scores. What's more, planning seems to be particularly helpful for the adverse conditions, namely the unfamiliar topics, and the lower proficiency levels.

5.8 Lexis

This section focuses on three lexical aspects in task performance. Since no interaction is found with these dependent variables, only the effects of the three independent variables will be outlined here in the following three tables.

Table 5.28 The effects of topic familiarity on lexis

	Topic familiarity			BG df	WG df		Cohen's *d*
	Familiar	Unfam.	F			Sig.	
D (Lexical diversity)	52.33 (11.17)	49.04 (11.87)	5.84	1	76	.018	.29
Lambda (Lexical Sophistication)	2.80 (.45)	2.62 (.46)	15.10	1	76	.000	.41
Lexical density	55.22 (3.73)	54.87 (4.22)	.56	1	76	ns	/

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.29 The effects of planning on lexis

	Planning		F	BG df	WG df	Sig.	Cohen's *d*
	Unplan'd	Planned					
D (Lexical diversity)	51.33 (12.55)	50.04 (10.51)	.34	1	76	ns	/
Lambda (Lexical Sophistication)	2.65 (.43)	2.77 (.47)	1.79	1	76	ns	/
Lexical density	54.07 (3.86)	56.02 (3.87)	7.33	1	76	.008	.43

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Table 5.30 The effects of proficiency on lexis

	Proficiency		F	BG df	WG df	Sig.	Cohen's *d*
	High	Intermed					
D (Lexical diversity)	52.48 (11.38)	48.89 (11.47)	2.66	1	76	ns	/
Lambda (Lexical Sophistication)	2.74 (.39)	2.67 (.51)	.62	1	76	ns	/
Lexical density	55.84 (3.59)	54.25 (4.20)	4.85	1	76	.031	.39

Note: 1. Standard deviations in (). 2. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

Topic familiarity is an important factor for lexical diversity, as indexed by the value of *D*, and lexical sophistication, as indexed by the value of *Lambda*, with a small and a medium effect size respectively, but familiarity does not influence lexical density ($p > .05$). Speaking on a more familiar topic, participants were able to draw on a larger repertoire of vocabulary and recycle the same set of words less ($p = .018$, $d = .29$). Also, the higher *Lambda* value ($p = .000$, $d = .41$) indicates that familiarity with a certain topic appears to be connected with a range of lexical items specialized in that field, which are rarer in use elsewhere. Planning has no effects on either lexical diversity or lexical sophistication, but it promotes a higher ratio of content words, thus higher lexical density (p = .008), as further evidenced by the medium effect size (d = .43). Similar to planning here but with a smaller effect size, higher proficiency is helpful for the use of more content words (p = .031, d = .39), but irrelevant to the other two measures. With no interaction found, the main effects of each independent variable are clear-cut.

Summary of lexis:

Topic familiarity drives learners to a higher degree of lexical diversity and lexical sophistication. By contrast, *planning* and *proficiency* favor higher lexical density.

5.9 Formality

Two formality measures, namely the F-score and DB-involved words, are reported in this section. Only the main effects will be included here due to the lack of interactions.

The F-score and the D. Biber's 'involved style' score are significantly correlated (r = -.42, p = .000 on the familiar topics, and r = -.48, p = .000 on the unfamiliar topics, N = 80 for each correlation), suggesting that the twomeasures may share some common ground as indices for how formally one speaks.

Table 5.31 The effects of topic familiarity on formality

	Topic familiarity		F	BG df	WG df	Sig.	Cohen's *d*
	Familiar	Unfam.					
F-score	84.71	74.66	16.89	1	76	.000	.49
	(25.66)	(16.20)					
DB-Involved	7.66	8.36	2.57	1	76	ns	/
	(2.86)	(3.76)					

Table 5.32 The effects of planning on formality

	planning		F	BG df	WG df	Sig.	Cohen's *d*
	Unplan'd	Planned					
F-score	73.69	85.67	9.23	1	76	.003	.48
	(18.97)	(21.09)					
DB-Involved	8.72	7.30	5.72	1	76	.019	.38
	(3.20)	(3.27)					

Table 5.33 The effects of proficiency on formality

	Proficiency		F	BG df	WG df	Sig.	Cohen's *d*
	High	Intermed					
F-score	80.40	78.96	.13	1	76	ns	/
	(21.17)	(20.82)					
DB-Score	7.74	8.28	.83	1	76	ns	/
	(3.13)	(3.50)					

Note for Table 5.31 -- 5.33: 1. F-score is calculated based on pruned speech. 2. DB-Involved is calculated by the 'involved' words per 100 words. 3. Standard Deviations in (). 4. BGdf= between-group degree of freedom, WGdf=within-group degree of freedom.

That being said, Table 5.31above shows that, though arithmetically both measures tell us that the participants spoke more formally on the more familiar topics, familiarity only makes a significant difference in the F-score with a medium effect size (p = .000, d = .49). Participants performing a familiar task do not necessarily use fewer 'involved' words (p > .05, and the lower the score, the more formal it is). Planning appears to be a stronger variable in promoting more formal language in that both measures reached significance with a similar effect size to that of topic familiarity. A chance to carry out pre-task planning prepares speakers for more 'nouny' language (F-score) (p = .003, d = .48) and a less 'involved' speaking style as the fewer 'involved' words suggest (p = .019, d = .38). Proficiency does not seem to be a factor in either of the measures. Nor does any interaction occur with the formality variables.

Summary of formality:

Both familiarity and planning push learners for higher F-scores, but only planning significantly reduces the 'involved' words, to be regarded as speaking more formally. Proficiency, however, does not exert any impact on the two measures.

The previous sections dealt with all dependent variables in a detached manner where results of each measure were presented separately, though some of them were considered different aspects of the same construct theoretically and grouped together into one cohort. It seems beneficial, at this point, to go beyond individual measures and take a top-down view at their interrelationship so as to facilitate our discussion in the next chapter.

5.10 Underlying constructs in L2 oral production

The rationale for the adoption of the range of measures reported above to assess performance lies in theoretical discussion as well as past empirical research. Beyond these claims we need to explore to what extent the data in the present study is consistent with theories or previous studies. Two factor analyses of the two task types (familiar vs unfamiliar) were conducted to probe into the constructs underlying L2 oral performance, and this section covers the results.

The appropriateness of the data for factor analysis was examined prior to the actual statistical procedure. First of all, many dependent variables in this study, as mentioned earlier in this chapter, are significantly related, especially those subsumed under the same general category, suggesting a converging tendency among some variables. Second, the Kaiser-Meyer-Oklin measure of sampling adequacy value is .64 for the familiar tasks, and .68 for the unfamiliar tasks, both exceeding the requirement of .60 proposed by Kaiser (1974). Barletts Test of Sphericity is also very significant, approx. x^2 (300)= 1720.67, p = .000 in the familiar condition, and approx. x^2 (300)=1843.46, p = .000, supporting the factorability of each correlation matrix.

The default solution adopted by the factor analysis module in SPSS 17 is to determine the number of factors based on those whose Eigenvalues are above 1.

This solution is adopted in this study out of three considerations. The primary concern is, of course, that the default method is mathematically the most reasonable solution. Then, with seven factors as the default number, other solutions (from four to eight factors) were tried, but none of the other numbers had loadings on better commonalties than the seven-factor solution. Lastly, though with different orders of the constructs, both familiar and unfamiliar tasks had loadings on exactly seven components, making a cross-task-type comparison potentially more meaningful and interesting.

Tables 5.34 and 5.35 show that the two task types display striking similarities in the factor structure, suggesting that the particpants behaved stably across task types. The first component is clearly a breakdown fluency factor that had very high loadings on mean length of run, speech rate, phonation time, numbers of pauses and silence total at both mid-clause and end-of-clause positions, on both familiar and unfamiliar tasks. This factor also loads on repetitions and words per AS units, though with loadings not as high as those distinctive temporal-acoustic aspects. The negative loadings confirm the obvious reasoning that those who spoke faster, with a longer smooth stretch of words and more speaking time would pause less frequently and stay silent less, with fewer repetitions. What remains less apparent though was the indication from this factor that if one is able to pack more clauses into one AS unit, the clauses are more likely to be produced as a chunk.

The second factor, also consistent across familiarity types, is a clear-cut accuracy dimension, with three different but related measures. The picture is

Table 5.34 Factor analysis of all dependent variables in the familiar tasks

	Component						
	1	2	3	4	5	6	7
Mean Length of Run	**-.849**	.015	.093	.060	-.309	.062	.027
Words per Minute	**-.862**	.105	-.211	-.072	-.262	.037	-.034
Phonation Time	**-.858**	.055	**-.382**	.060	.211	.074	.014
Mid-Clause Pause Length	.463	-.062	**.751**	.006	-.114	-.097	.073
End-of-Clause Pause Length	.425	-.033	**.767**	.116	-.129	-.063	.004
Mid-Clause Pause Number	**.923**	-.056	.070	.027	.032	.044	.036
End-of-Clause Pause Number	**.856**	-.107	-.146	-.140	-.096	.034	-.145
Mid-Clause Silence Total	**.834**	-.108	**.404**	.007	-.024	.014	.078
End-of-Clause Silence Total	**.833**	-.129	**.382**	-.037	-.149	-.030	-.113
Filled Pause Number	.443	.012	-.165	.160	**.601**	-.110	.089
False Starts	-.038	-.040	**.685**	-.055	.125	.099	-.282
Reformulations	.131	-.074	.141	-.264	**.724**	.029	-.157
Replacements	-.181	-.006	-.073	.046	**.650**	-.021	.034
Repetitions	**.367**	-.081	.264	-.296	.271	-.085	.307
Error-free clause ratio	-.109	**.920**	-.083	-.019	-.117	.098	.099
70% Accuracy Clause length	-.047	**.828**	.014	.203	.147	.052	.014
Errors per 100 words	.174	**-.888**	.077	-.203	.121	-.029	-.108
Clauses per AS unit	-.023	.142	-.107	-.134	-.013	.098	**.892**
Words per AS unit	**-.455**	.246	-.333	.304	-.113	-.074	**.523**
Words per Clause	-.097	.243	.110	**.747**	.042	-.164	-.186
D (Lexical diversity)	-.284	.083	.204	-.153	-.089	**.711**	-.076
Lambda (Lexical Sophistication)	.008	.065	-.034	**.536**	-.009	**.481**	.123
Lexical density	.211	.094	-.223	.301	.020	**.747**	.106
F-Score	-.062	.220	.015	**.756**	-.032	-.033	-.088
DBInvolved	-.060	.083	.062	**-.749**	.062	-.260	-.101

Note: 1. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. 2. Rotation converged in 11 iterations. 3. All pause numbers, silence total and repair measures were standardized by calculating their occurrence per 100 words.

Table 5.35 Factor analysis of all dependent variables in the unfamiliar tasks

	Component						
	1	2	3	4	5	6	7
Mean Length of Run	**-.824**	.191	.139	.089	-.091	.008	-.152
Words per Minute	**-.899**	.112	.063	-.172	.021	.103	-.175
Phonation Time	**-.791**	.128	.046	**-.462**	.244	.008	-.060
Mid-Clause Pause Length	**.738**	.003	.081	.335	-.281	-.081	.194
End-of-Clause Pause Length	.235	-.002	-.033	**.874**	-.068	-.062	.098
Mid-Clause Pause Number	**.893**	-.084	-.016	.118	-.139	-.040	.126
End-of-Clause Pause Number	**.737**	-.258	-.186	.239	-.045	.239	-.088
Mid-Clause Silence Total	**.841**	-.006	.074	.239	-.218	-.027	.187
End-of-Clause Silence Total	**.625**	-.114	-.100	**.685**	-.102	.132	.023
Filled Pause Number	**.750**	-.002	.059	-.326	.217	-.089	.123
False Starts	.312	.049	-.003	.054	-.352	.006	**.538**
Reformulations	.194	-.040	-.052	.061	.167	.133	**.745**
Replacements	.090	-.271	.262	-.146	**.539**	-.049	.231
Repetitions	**.492**	-.092	.028	.120	.156	-.232	**.518**
Error-free clause ratio	-.066	**.900**	-.229	-.050	.091	.055	-.029
70% Accuracy Clause length	-.101	**.846**	-.064	.032	-.008	-.023	-.004
Errors per 100 words	.210	**-.876**	-.108	.079	-.099	-.036	.033
Clauses per AS unit	-.105	.309	-.157	.026	**.782**	.217	-.068
Words per AS unit	**-.500**	.219	.171	-.193	**.655**	-.004	.021
Words per Clause	-.121	.061	**.822**	-.111	-.116	-.200	.149
D (Lexical diversity)	-.135	.144	-.299	-.103	-.089	**.665**	.350
Lambda (Lexical Sophistication)	-.141	.207	**.450**	.061	.141	**.462**	-.415
Lexical density	.025	-.063	.135	.037	.149	**.898**	-.064
F-Score	-.137	-.131	**.808**	-.140	.126	.028	-.047
DBInvolved	-.171	.197	**-.731**	-.229	-.054	-.229	.219

Note: 1. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. 2. Rotation converged in 9 iterations. 3. All pause numbers, silence total and repair measures were standardized by calculating their occurrence per 100 words.

remarkably straightforward in that while the three variables have loadings close to .90, few of the other variables are above .20.

The third factor from the familiar tasks echoes the fourth dimension in the unfamiliar tasks, which appear to be a fluency measure different from the clear breakdown fluency factor or the repair fluency factor that will be discussed below. This factor has high loadings on the mid-clause pause length, end-of clause pause length and false starts, with smaller loadings on mid-clause and end-of clause silence, for the familiar conditions. Meanwhile, situations on the unfamiliar tasks are simpler in which only the pause length and total silence at the clause boundaries enjoy high loadings, with a medium loading on phonation time. This factor is not an entirely unambiguous construct, but two aspects do seem to stand out. Firstly, pause length and silence total between clauses seem to be less connected with the breakdown fluency measures, which is consistent across two tasks, and confirms the result in Skehan and Foster (2005). Secondly, pause length appears to distance itself from the number of pauses, and even other fluency measures.

The fourth factor emerges neatly again with striking resemblance in both task types, though this fourth factor is ordered the third in the unfamiliar task. High loadings are associated with the formality indices, the F-score and DB-Involved, with an interesting and equally high loading in words per clause, which was supposed to be a complexity index. This factor also loads slightly less but significantly on lexical sophistication. This structure indicate that the more 'nouny' the language, the less 'involved' one would be in speaking, with greater clause length, and probably more infrequent words.

The fifth factor in the familiar tasks could be regarded as the same as the seventh factor in the unfamiliar, with clear indication of a repair fluency construct. The familiar tasks see reformulations, replacements and filled pauses in one factor, whereas the unfamiliar tasks include false starts, reformulations and repetitions. The sixth factor in both task types is the lexical construct, with an identical pattern and even the same order in the components. An intriguing point here is that lexical sophistication seems to be less associated (but still significantly) with this construct while lexical diversity and lexical density are more central, or rather, reflect more on the same construct.

The last construct in the familiar tasks could be regarded as the same as the fifth dimension in the unfamiliar task, both concerning the complexity with which one uses language. Clauses per AS unit and words per AS unit are closely related, as it is reasonably for an AS unit to contain more words if there are more clauses in one unit. Both tasks suggest that the newly developed 'words per clause' does not appear to hold a legitimate status in this complexity construct.

Summing up, this study largely confirms the constructs outlined at the literature and methodological sections as the general categories of measures, but also with some new findings. The seven factors seem to be: breakdown fluency, repair fluency, end-of-clause fluency, accuracy, complexity, lexis, and formality. Table 5.36 is a summary of constructs that are consistent in both tasks (Table 5.34

and 5.35). For example, 'repetitions per 100 words' appeared in 'breakdown fluency' in the familiar task, but in both 'breakdown fluency' and 'repair fluency' in the unfamiliar task. Then 'repetition' will, though counter-intuitively, be put under the breakdown fluency construct only. Two levels of loadings are distinguished: above .50 (strong) and between .35 – .50 (moderate), so as to further clarify the strength of association between variables. As could be noted in Table 5.36, some construct names have undergone modification compared to the 7 constructs mentioned earlier, which is the result of the analyses on some intriguing 'deviations' from the literature.

Firstly, both tasks indentified a third fluency construct closely related to measures for fluency at the end of a clause, which could indicate online processing and probably a certain degree of native-like proficiency. Second, repair fluency seems rather unstable as it appears to share some common ground with breakdown fluency sometimes. Thirdly, the 'words per clause' measure was found to be linked to the formality measures and seems irrelevant to complexity. Fourthly, *Lambda* (lexical sophistication) seems to be split between both lexical aspects and formality, but this should not surprise us if people use more rare words in formal speech. Combining the third and the fourth points, this 'formality' construct emerged more as noun phrase complexity index with all relevant measures closely associated with 'nouny' language and the length and quality of noun phrases. Lastly, words per AS units, which was supposed to be a complexity measure, also appears to be divided between breakdown fluency and complexity. All these

Table 5.36 Constructs consistent across task types

	Measures	
	Loadings above 0.50	Loadings between .35 and .50
Constructs		
Breakdown fluency	1. mean length of run 2. speech rate 3. phonation time 4. mid-clause pause number 5. end-of-clause number, 6. mid-clause silence, 7. end-of-clause silence	1. mid-clause pause length 2. filled pauses 3. repetitions 4. words per AS unit,
End-of-clause (Online) fluency	1. end-of-clause pause length	1. phonation time 2. end-of-clause silence
Repair fluency	1. reformulations	
Accuracy	1. error-free clause ratio 2. 70% accuracy clause length 3. errors per 100 words	
Complexity	1. clauses per AS unit 2. words per AS unit	
Lexis	1. lexical diversity 2. lexical density	1.lexical sophistication
Noun phrase complexity	1. words per clause 2. F-score 3. DB-score	1. lexical sophistication

convergences and divergences from past studies will be further explored and theorized in the subsection 6.5.5 in Chapter 6 that follows.

These two factor analyses should also have methodological implcations for future selection of measures rèpresenting each general construct. Such determination is obtained though the loadings for each measure in the components. The number of mid-clause pauses (per 100 words), speech rate, and mean length of run are the most significant measures for the construct of "breakdown fluency" and they functioned consistently across task types. Reformulation seems to be the only stable measure for "repaire fluency" in both factor analyses. For "accuracy", though the three measures are highly correlated, this construct is most loaded in the classic measure of "ratio of error-free clauses to all classes". Similar situation occurs to "complexity", as the classic "clauses per AS unit" stands out from both task types. Both the "F-score" and the"words per AS unit" measures appear to lie in the heart of the newly discovered construct "noun phrase complexity". The three lexical measures are considered different constructs and it would seem inappropriate to affirm which is the most representative. They should be treated as individual measures in their own right.

5.11 Relations between accuracy and complexity

The purpose for the inclusion of this section is to provide a detailed description of the relations between accuracy and complexity so as to find clues for the well-known Robinson-Skehan debate in TBLT research. Though in Section 5.10 we have seen that accuracy and complexity are separate constructs, it is not so clear how their relations vary in response to proficiency levels as we might

Table 5.37 Correlations between accuracy and complexity

Proficiency			C-test	Accuracy	Complexity 1
Familiar Tasks	High	Accuracy	.340*		
		Complexity1	.070	.062	
		Complexity 2	.387*	.203	.479**
	Inter-mediate	Accuracy	.390*		
		Complexity1	-.186	.249	
		Complexity 2	.134	.233	.451**
Unfamiliar Tasks	High	Accuracy	.480**		
		Complexity1	.261	.426**	
		Complexity 2	.423**	.238	.468**
	Inter-mediate	Accuracy	.361*		
		Complexity1	.027	.142	
		Complexity 2	.094	.081	.719**

Note: 1.Accurayc= ratio of correct clauses to all clauses. Complexity 1= clauses per AS unit. Compexity 2= words per AS unit. 2. N = 80

surmise that the trade-off effect could probably be resolved when one achieves a higher level of automatization in L2. Table 5.37 above shows their relations at different proficiency levels in each task types.

The two task types show almost identical patterns except that participants of higher proficiency levels were able to achieve better accuracy and complexity at the same time in the unfamiliar topics (accuracy correlates with complexity (clauses per AS unit), $r = .426, p < .01$, n = 80) but not in the familiar topics. What appears consistent in both tasks are: 1) accuracy is always significantly correlated with the proficiency test, regardless of task types or proficiency levels (r ranges from .340 to .480, $p < .05$ in all four conditions); 2) accuracy and complexity (clauses per AS unit) are not correlated in all but one case mentioned above; 3) Complexity (length of AS unit) is correlated to proficiency at the high proficiency level ($r = .387, p < .05$ in the familiar topics, and $r = .423, p < .01$ in the unfamiliar topics), but not the intermediate; 4) the two complexity measures were always very significantly correlated (r ranges from .451 to .719, $p < .01$ in all four conditions).

Chapter 6 Discussion

The preceding chapter presented results on each dependent variable and their interrelationship. This chapter will further discuss the results in terms of the independent variables and their interrelationship so that the effects of topic familiarity, planning and proficiency can be explored more directly. Before the discussion for each independent variable, a general framework of task readiness will be dealt with as the basis for the subsequent sections.

6.1 A general framework of task readiness

Ellis (2005) distinguished between two types of planning: 1) pre-task planning which can be further divided into rehearsal and strategic planning, and 2) within-task planning that is further split into a pressured and an unpressured situation according to the time allowed during performance (see Figure 1 in Chapter 2). Ellis (in press) slightly revised this set of categorization and talks about three types of planning: rehearsal, pre-task planning and within-task planning. These two categorizations are essentially the same, all dwelling on task-external manipulations of the degree of preparation for a task. Rehearsal and pre-task (strategic) planning without doubt prepare learners prior to a task, but within-task planning can also be viewed as increasing the readiness for performance in a series of consecutive segments of strategic planning, carried out in bits and pieces during a task.

What seems interesting is, if we adopt a broader perspective on this issue, that the notion of planning as preparations or readiness in order to do a task (better) should extend its horizons beyond these task-external means outlined in Ellis (2005, in press). One's prior knowledge about, and hence familiarity with, the content of a task or the schemata of a task should also be drawn into a broader sense of planning. This study has provided evidence that familiarity with a certain topic facilitates learner performance in a variety of ways similar to other types of planning, albeit different in some other areas as well. Therefore topic familiarity is, one could argue, a kind of task-internal readiness, or implicit planning, as contrasted with task external readiness, or explicit planning, with the latter reviewed in Ellis (2005, in press). The following table displays this extension of the construct of planning.

Table 6.1 A framework of learner readiness for a task

	Macro-dimension	Micro-dimension	Sample studies
Learner readiness for a task	• **Task-internal** readiness (implicit planning)	• **Topic familiarity** (prior domain knowledge)	This study
		• **Schematic familiarity** (story structure)	Skehan and Foster (1999)
		• **Task familiarity** (task types)	Bygate (2001)
	• **Task-external** readiness (explicit planning)	• **Rehearsal** (content repetition)	Bygate (1996)
		• **Strategic** (pre-task) planning	Foster and Skehan (1996)
		• **Within-task** (on-line) planning	Yuan and Ellis (2003)

As shown in Table 6.1, learner readiness consists of two macro-dimensions. Obviously what the second macro-dimension involves are the three common planning types, those discussed in Ellis (2005, in press). The novel part here is the first macro-dimension, the task-internal readiness, which also subsumes three different aspects.

The first kind of task-internal readiness is topic familiarity, which is the prior knowledge about a certain domain area, such as medical knowledge on a natural virus by a medicine major, or the technical specialty about computer virus by a computer major, as exemplified in the present study. The second kind concerns schematic familiarity. Examples can be found in Skehan and Foster (1999) in which a 'going to a restaurant' in the Mr. Bean video stood out as a fairly predictable story because nearly everyone has a schema of 'coming in – order the dishes – eat the meal – pay the bill – leave the restaurant' in mind. Compared with the more predictable storyline in a restaurant, what happened when Mr. Bean played golf was hard to foresee due to the lack of the relevant schema. The third type of task-internal readiness is task familiarity which deals with whether there will be a practice effect transferred from a task to another of the same type (but different in topic), with Bygate (2001) as a case in point. Our subsequent discussion on the effects of topic familiarity and planning will be based on this general framework of task readiness.

The biggest difference between task-internal and task-external readiness is the degree of naturalness, or rather the degree of *ad hoc* manipulation, of the task preparation. Task-internal readiness, especially topic familiarity and schematic familiarity, could be thought of as a more inherent and natural way of planning, albeit perhaps not so much a conscious process. At the same time, task-external means are more on an artificial side where learners are imposed upon them extra manipulations to a task, with of courşe usually favorable conditions. A question then arises from this comparison: which is a stronger driving force for the improvement of task performance? The following sections, 6.2 and 6.3, will explore this task-internal and task-external dichotomy in task performance, taking topic familiarity and strategic planning as relevant instances. Section 6.4 furthers the discussion by going through the effects of proficiency and its interactions with topic familiarity and planning, followed by section 6.5 towards a more complete picture for this general framework.

6.2 Topic familiarity

Research question 1 'what are the effects that topic familiarity exerts on L2 oral performance' has been answered in Chapter 5 'Results of main study'. Briefly, topic familiarity was able to push learners for more total words, greater fluency with fewer breakdowns, more diverse and more sophisticated lexis, more 'nouny' language, and slightly higher accuracy and repair fluency. What topic familiarity was not so effective in were complexity, end-of-clause fluency, lexical density and the degree of an 'involved' style. Therefore, Hypotheses 1, 2, 3 on the main effects of topic familiarity are mostly confirmed, except that topic familiarity did not show its influence on lexical density in the predicted way.

Several indications from the results are note-worthy: firstly, topic familiarity seems to affect both the conceptualization and the formulation stages in Levelt's (1989) speaking model. The Conceptualizer is responsible for drawing information from memory and forming a pre-verbal message for the formulation stage next. It will take less time to access the more familiar information due to the immediacy effect, since speakers are more primed about the relevant knowledge domain. As a conceptualizer effect, too, speakers have a more ready-made schematic structure at their disposal which could be accessed on a macro basis. As shown in the results concerning 'total words' (360.36 raw words and 300.84 pruned words as compared to 284.05 raw words and 229.61 pruned words in the unfamiliar topics, p = .000 for both topics), the speedily accessible message plus an existing framework into which the message can be structured can ease the workload at the conceptualization stage. The longer account produced in the familiar topics indicates that more familiar information can be retrieved from the long term memory in any given time period.

The significantly higher *Lambda* (lexical sophistication) value in the more familiar topic (*Lambada* = 2.80 and 2.62 for the familiar and unfamiliar topics respectively, p = .000, d = .41) further enhances the argument that topic familiarity has an effect on conceptualization. It is natural that learners with a certain specialty are in possession of a set of terminology specific to their major, which has a much lower frequency of use outside of this register. Medicine majors could only use general vocabulary to describe the (unfamiliar) computer virus while computer majors were much better at pulling in professional terms on their (familiar) computer virus task, and visa versa. As Skehan (2009) pointed out, lexical sophistication has a closer connection to the conceptualization stage of the Levelt model, and to the nature of preverbal message implications for lemma retrieval. It could be inferred from this result that the more familiar topic helps to activate a more organized and larger part of mental lexicon for speaking.

Topic familiarity also appears to exert an influence on the articulator stage in Levelt's (1989) model. The articulator receives the non-linguistic information from the conceptulizor, then draws on lemmas and lexemes from the mental lexicon and assembles them into a linguistic plan waiting to be articulated at the next stage. In this process, lexis can be retrieved not only at higher speed (with fewer breakdowns as the proof, see Table 5.6), but also in a larger quantity, as especially evidenced by the higher *D* value (52.33 in familiar topics Vs. 49.04 in unfamiliar topics, p = .018, d = .29) to suggest less recycling of the same set of vocabulary items, when speaking on a more familiar topic. If lexical diversity (*D*) means quantity, a larger pool of different lexis, lexical sophistication may then suggest quality, a store of infrequent words to choose from.

Aside from the individual word choice, longer mean length of run (5.26 for familiar Vs 4.99 for unfamiliar topics, p = .016, d = .17), higher phonation time (0.80 for familiar Vs 0.77 for unfamiliar topics, p = .000, d = .35) and fewer mid-clause pauses (9.73 for familiar Vs 12.13 for unfamiliar topics, p = .000, d = .38) all suggest that the more familiar topics are able to promote bigger chunks in which more lexical items are packed into an uninterrupted utterance, which is an indication that topic familiarity helps learners with lexicalized language. This would not only explain the better temporal aspects of speaking, but also the slightly, but significantly higher accuracy results because if some expressions are memorized as a whole, it reduces the analytic workload and thus error probability. Also, the higher F-score (84.71 for familiar Vs 74.66 for unfamiliar topics, p = .000, d = .49) as a formality index in the more familiar situation provides further evidence of the book-based (familiar) speech than the casual (unfamiliar) speech. To sum up, learners were able to more efficiently access the exemplar-based system (Skehan, 1998) when they are in possession of the relevant prior knowledge.

An additional explanation that might not be as general as those discussed above may nonetheless apply well to this context specifically. The medium of instruction for the major courses in both academic groups is primarily English, and all the textbooks and lecture notes are in English. According to the encoding specificity principle (Tulving and Thomson, 1973), the language in which knowledge is stored in long term memory will speed up access. Therefore, participants might have to go through one more step in the formulation stage, that of transforming their general knowledge about the unfamiliar topic from Chinese into English, which certainly hampers their performance in terms of fluency and lexis. This might have some implications on content-based language teaching in that if a certain domain knowledge is learnt in one's L2, it appears that future retrieval of the knowledge and production in L2 will benefit at least as far as fluency and lexis are concerned (see the 'pedagogical implication' section below).

More generally speaking, under the un-planned condition, all this occurred as pressured on-line planning (Ellis, 2005). Their limited processing capacity (Skehan, 1996) creates difficulties for L2 speakers whose target language system is not yet automatized to do parallel processing and more attentional resources allocated to the conceptualization stage mean difficulties in the later formulation and articulation stages. Therefore, learners had to slow down their speech rate and pause more often with a shorter average speaking time in order to cope with the unfamiliar topics. This result for fluency is largely consistent with some studies in L1 (e.g., Good and Butterworth, 1980; Bortfeld et al., 2001) and L2 (Chang, 1999; Skehan and Foster, 1999; Robinson, 2001). However, pre-task planning is able to attenuate the difference between unfamiliar and familiar topics in many of the fluency measures, and this will be discussed in sections 6.3 and 6.5 below.

The second point to consider is the form-meaning connection in relation to topic familiarity. The primary concern on the part of learners in a speaking task is obviously to get the message across. Meaning expression is more likely to be attended to than the other aspects of speaking. However, it appears that the familiar topics also raise accuracy (ratio of error-free clauses: .544 for familiar Vs .517 for unfamiliar topics, $p = .020$, $d = .22$), achieving meaning and form at the same time.

In addition to the theory of better access to the exemplar system and chunking, two more possibilities from a processing perspective are available. First of all, the attentional resources released from the conceptualization and the articulation stages can help learners with self-monitoring. In the more familiar topics, speakers may shift their attention focus partly from 'what to say' to 'how to say' and even 'how to say well', whereas they will have to struggle with the content to express in the unfamiliar situations, which results in less working memory load for monitoring and correction. Secondly, on-line planning studies (e.g., Yuan and Ellis, 2003) provided evidence that unpressured within-task planning can contribute to more accurate performance. As a task-internal readiness construct, topic familiarity appears to achieve a similar effect because it prepares learners not only prior to the task, but through the whole process of speaking. It is plausible that this on-line readiness resemblance to unpressured on-line planning may partly explain the higher accuracy scores in the familiar tasks. At the same time, the small effect sizes (.22 for error-free ratio and .38 for errors per 100 words) may as well be justified simply because the task-internal readiness is still time-pressured when compared to the unpressured task-external on-line planning.

6.3 Strategic planning

This section discusses research question 2: 'What are the effects that strategic planning exerts on L2 oral task performance?' A plethora of studies have investigated the effects of strategic planning in task-based language performance, with quite consistent results in fluency and complexity, some convergent results in lexical aspects (although this performance area has not been included in many studies), and mixed results in accuracy (see Ellis, 2005 and in press, for comprehensive reviews). The present study has confirmed the literature very well in fluency (hypothesis 4), largely in complexity and accuracy (hypothesis 5), but not in lexis. Interesting results in formality will be discussed as well though there is almost no study reporting on this construct.

Judging from the range of measures planning has effects on, and the effect sizes that it produces, planning stands out as a more powerful means in improving fluency than topic familiarity. That is, this task-external planning constitutes a higher level of task-readiness than the task-inherent planning as far as fluency is concerned. The ten-minute planning time allows learners to formulate a conceptual plan for the relevant message to convey rather than a detailed linguistic plan (Ellis, 2005). Planning does not seem to change learners' preference from meaning to form as their primary concern. A comparison of the results shows that planning works in a pretty similar way to topic familiarity, though obviously planning is more effective, especially in reducing average pause length and repairs. Therefore similar explanations for the effects of fluency will not be repeated here.

What distinguishes topic familiarity from planning is that planning pushes learners to higher complexity ('clauses per AS unit': 1.81 for planners Vs 1.67 for non-planners, p = .018, d = .39, and 'word per AS unit': 13.96 for planners Vs

11.39 for non-planners, $p = .000$, $d = .81$). The lexicalized language, or the chunks, that are more readily and speedily accessible due to topic familiarity cannot be very complex syntactically as we know intuitively that the longer a chunk is, the harder it is to remember. Therefore, a reasonable assumption here would be that those available in long term memory are usually relatively short expressions. A comparison with topic familiarity may show that strategic planning helps learners not only to access formulaic language (Foster, 2001[7]) and hence achieve higher fluency, but also assemble the pre-fabricated chunks into a longer psychological unit of planning (AS unit), as shown in more words per AS unit and higher subordination ratio. This result is consistent with most studies, that planning drives learners to take risks for more elaborated language. To some extent, this study, combined with Foster (2001), helps to better explain why task-external readiness can, but task-internal readiness cannot, promote complexity

Rather disappointingly, strategic planning does not seem to affect the measure of clause length 'words per clause' even though it is supposed to be a complexity measure for more advanced learners as argued in Ortega, Iwashita, Norris, and Rabie (in preparation). Two possibilities exist here: first, planning cannot promote complexity and second, 'words per clause' is a disguised measure that reflects a different construct other than complexity. If we accept that the commonly employed subordination measure and the AS length measure are

[7] Foster (2001) found that, given planning time, native speakers tend to use less formulaic language and be more creative, whereas non-native speakers will use more formulaic language after planning.

genuine complexity indices, we may rule out the first possibility. Furthermore, the two factor analyses in section 5.10 both confirmed that 'words per clause' appears to be very closely connected to the F-score and the DB-Involved score, and less closely but significantly with lexical sophistication. As will be further discussed in subsection 6.5.5 below, this construct could be regarded as a noun phrase complexity, as distinctive from the syntactic or lexical complexity indentified in the literature.

What remains opaque is the relationship between planning and accuracy. The past literature has been mixed in this respect, and the present study did not find a significant accuracy effect from planning. A thorny question emerges naturally at this point: if as mentioned above, planning enables L2 learners to better access their lexicalized language (formulaic chunks) as topic familiarity does, why can topic familiarity raise accuracy but planning cannot? Possibly the puzzle can be solved through three proposals. Firstly, planning drives learners to embark on more complex language and in the process more pre-fabricated expressions need to be assembled into an AS unit. The more syntactic work there is, the more errors there might be (Crookes, 1989). Secondly, from a limited processing capacity point of view (Skehan, 1996), there is likely to be trade-off between accuracy and complexity (Skehan and Foster, 1997a). Learners' L2 systems are, by and large, a controlled but not an automatized one, and so attentional resources allocated to the overwhelming workload of complexity build-up means a reduction of attentional focus on accuracy. Thirdly, it is possible that pre-task planning cannot affect on-line monitoring (Skehan, 2009), but topic familiarity as a task-inherent readiness prepares learners anytime they speak, acting as both pre-task and on-line readiness, and reduces the on-line processing workload to enable more within-task monitoring.

Turning to the lexical effects, there are no differences between planners and non-planners in terms of lexical diversity or lexical sophistication ($p > .05$ for both), but planning does promote higher lexical density (56.02 for planners Vs 54.07 for non-planners, $p = .008$, $d = .43$). Many studies (Wendel, 1997; Ortega, 1999; Tajima, 2003; Yuan and Ellis, 2003) did not find a planning effect on lexical complexity, based on which Ellis (in press) concluded that the effect of planning on lexis is marginal. Skehan (2009) found that planning did not influence lexical diversity, but in personal and decision-making (but not narrative) tasks, lexical sophistication can be elevated by planning.

The present study is consistent with the literature that planning is irrelevant to lexical diversity, as indexed by the corrected type-token ratio, the *D* value ($p > .05$, there is no difference between planners and non-planners in *D*). It appears that both video-based narratives (Skehan, 2009) and topic-based narratives (this study) are resistant to the effects of planning on *Lambda*, the lexical sophistication measure. Lexical diversity seems to be more concerned with the formulation stage, as it behaves as an on-line processing construct involving minute-by-minute decision (Skehan, 2009). Therefore pre-task planning could hardly have an impact on *D*.

The fact that *Lambda*, as seemingly a conceptualization variable, is not affected by planning is curious (2.77 for planners Vs 2.65 for non-planners, $p > .05$). Taking this study and the several studies motioned in Skehan (2009) together, it appears that this result is at least partly a task effect. In a topic-based narrative task, neither input nor interlocutor scaffolding is available. What learners can resort to during planning is only the mental lexicon in their longer memory. Compared to the set of infrequent words related to their academic endeavor, the chance to draw on the same amount of rare words for the unfamiliar topic is slim because there are not many in stock. However, the higher lexical density does suggest that planning can help learners to pack more content words into speech. This is interesting in that we might first of all think of the connection between more content words and more information load in the performance. If this holds true, it is striking to see that planning induces a greater information load which is expressed neither in more varied nor in more difficult words.

Interestingly, the correlations shown in table 6.2 below reveal that higher lexical density does imply both higher lexical diversity and sophistication (lexical density correlates with lexical diversity, $r = .324$ and .486 for the familiar and unfamiliar topics respectively, $p < .01$; lexical density correlates with lexical sophistication, $r = .334$ and .490, $p < .01$), though *D* and *Lambda* are not correlated at all. It is fair then to say at this point that if learners use more content words, the words tend to be more diverse and less frequent. It appears that this is especially true under the unfamiliar situations, as indicated by the higher correlation

Table 6.2 Correlations between lexical measures

	D _F	*Lambda*_F	Density_F	D _UF	*Lambda*_UF	Density_UF
D _F	1	.107	.324**	.461**	.207	.167
*Lambda*_F		1	.334**	-.076	.534**	.186
Density_F			1	.311**	.280*	.468**
D _UF				1	-.050	.400**
*Lambda*_UF					1	.490**
Density_UF						1

Notes: 1) N=80. 2) _F = familiar task, _UF=unfamiliar task. 3) * $p<.05$, ** $p<.01$

coefficient (r) values, probably due to the smaller lexical pool available and hence a bigger proportion of unavoidable words.

The last category in the discussion of planning is a relatively new area, that of formality. Planners have both higher F-scores (85.67 for planners Vs 73.69 for non-planners, p = .003, d = .48) and DB-scores (7.30 for planners Vs 8.72 for non-planners, p = .019, d = .38), which suggests more 'nouny' language use and a less 'involved' style. Planning strategically then appears to affect a more careful and less personal style of speaking. This sheds light on the conclusion that both task-internal and task-external readiness directs learner language in a more 'bookish' way. Seeing deeper into the issue, formality appears to belong to the conceptualization stage where learners assess the situation and decide on the appropriate style to attend to. This result indicates that, after conceptualization, learners while planning select more "non-deictic" words, assembled in a more context-independent way at the formulation stage. That is, planning allows learners to re-organize their language from a simply structure such as 'N+V+N' into a more elaborated one like 'Art.+Adj +N+V+N+Prep+N', with most of the newly added elements within a noun phrase structure. Though planning does not significantly raised the average clause length, it probably can drive learners towards a bigger noun phrase structure and this potentially leads to higher noun phrase complexity.

6.4 Proficiency

Research question 3 concerns the effects that proficiency has on the various aspects of task performance, and this section discusses the results relevant to proficiency. The past task-based literature has not seen proficiency taken seriously by most researchers. The few exceptions (e.g., Wigglesworth, 1997; Kawauchi, 2005), however, did suggest that task performance as influenced by strategic planning differs according to learners' proficiency levels. The present study re-examines the effects of planning at different proficiency levels, whilst adding to it a new dimension of planning: topic familiarity.

In terms of the main effects, proficiency shows consistently strong effects on all accuracy measures (e.g., ratio of error-free clauses: .586 for high Vs .475 for intermediate learners, p = .000, d = .69) and some effects on complexity (words per AS unit: 13.49 for high Vs 11.85 for intermediate learners, p = .000, d = .52; p = .067 for the conventional clauses per AS unit measure), with performances of learners at the higher proficiency level being more accurate and more complex. These more advanced learners were also able to reduce the number of pauses between clauses (number of end-of clause pauses: 6.15 for high Vs 7.22 for intermediate learners, $p = .027$, $d = .48$) and pack more content words into their speech (lexical density: 55.84 for high Vs 54.25 for intermediate learners, $p = .031$, $d = .39$). However, proficiency seems to be, at least in this context, largely irrelevant to fluency (either breakdown fluency or repair fluency), lexis, formality, and even noun phrase complexity. Therefore, Hypothesis 6 that proficiency will show broad effects on L2 speaking tasks received only partial support. An emerging pattern from these results is that proficiency tends to have much greater influence on syntactic than semantic aspects of performance.

Learners of higher proficiency consistently made fewer errors in performance than their lower proficiency counterparts did, regardless of familiarity types or planning time. Furthermore, the "70% accuracy clause length" measure indicates that the lower error rate was not ac' eved by the avoidance strategy with which one might make fewer errors by resorting to shorter and simpler utterances. Higher proficiency participants in fact spoke with longer error-free clauses than the lower proficiency participants did. All this reveals that accuracy performance is basically a by-product of one's underlying linguistic competence. Correlations in Table 5.37 in Section 5.11 clearly provided evidence that accuracy is very closely linked to one's proficiency level, as the accuracy measure (ratio of error-free clauses) always significantly correlates with the C-test scores in both familiar and unfamiliar task types. The lack of interaction between proficiency and the other two variables (planning and familiarity) further supports this claim, and might partly explain why accuracy performance was less sensitive to task manipulations like strategic planning.

It could be argued that better performance in accuracy originates from two sources: a well-developed linguistic system and a good ability to monitor speaking. A more advanced linguistic system plays a main role with error-free utterances and it almost becomes a cliché to mention that the actual 'performance' is a reflection of implicit 'competence'. A more fully-fledged underlying system is usually a more automatized one, which frees up more attentional resources for monitoring errors. All this contributes to the significantly and consistently better accuracy performance among the higher level learners in all three accuracy measures. The medium to large effect sizes (Cohen' *d* values ranging from .57 to .77) suggest that the difference in accuracy between the two proficiency levels is substantial.

Only one, namely 'words per AS unit', out of the three complexity measures were significantly affected by proficiency. However, the effects of proficiency nearly reached significance in the conventional 'clauses per AS unit' measure (p= .067). These results suggest that proficiency did show its influence on syntactic complexity, though its effects were not as big as those for accuracy. That being said, it is an interesting phenomenon that the subordination complexity measure (clauses per AS unit) were not correlated with either proficiency at all or accuracy in most cases, in table 5.37, Section 5.11. This triangular relationship between accuracy, complexity and proficiency seems to indicate that though higher proficiency generally means better accuracy and complexity performance, each individual learner still has to choose one aspects to focus on, which lends support to the trade-off theory (Skehan, 1998 and elsewhere).

Compared to strategic planning, proficiency is much less a driving force for higher complexity; compared to topic familiarity, proficiency is a much more important indicator for higher accuracy. Therefore, we might postulate that L2 learners tend to opt for a conservative stance in speaking and try to avoid mistakes. Planning time encourages them to be more willing to task risks and use more elaborated language. Higher proficiency itself can liberate L2 learners from their timidity only to a limited extent. Taking all the above discussion about proficiency together, Hypothesis 10 and 11 appears to be largely supported except for the prediction that lexical aspects of performance will be strongely affected by proficiency.

The one and only significant correlation between accuracy and complexity in Table 5.37 (Section 5.11) may be worthy of our attention. It shows that the two measures can go with each other among learners of high proficiency with unfamiliar topics. This probably suggests that high proficiency learners can sometimes escape from trade-off effects, which might be a sign of native-like proficiency. However, a similar correlation was not observed for the familiar topics among the same group of students, which undermines our confidence in the 'high proficiency, no trade-off' hypothesis. If this significant correlation is not a pure coincidence, then this happening may indicate a rather unstable ability to focus on more than one aspect of performance.

The Table 5.37 correlations could perhaps shed some light on this issue. Participants in past studies were learners with a wide range of proficiency levels (Crookes, 1989), pre-intermediate learners (Foster and Skehan, 1996), early intermediate learners (Mehnert, 1998), post-beginners and intermediate learners (Ellis, 1987), intermediate learners (Ortega, 1995). Very few (e.g., Kawauchi, 2005; Wigglesworth, 1997) had learners of higher proficiency. The trade-off effect has been found among low to lower-intermediate learners. This study, though still seeing some trade-off between accuracy and complexity, seems to indicate a trend towards a more balanced focus on performance areas. First, none of the correlation coefficients were negative, suggesting that though the two cannot be raised at the same time, they do not necessarily repel each other. Secondly, the high proficiency learners did have accuracy and complexity significantly correlated in the unfamiliar situation. Given the higher proficiency levels in this study compared to most past studies, it seems that there exists a trend for the more advanced learners to reach a higher level of automatization and thus suffer less from the limited processing capacity.

Regarding meaning expression, proficiency has only an effect on two measures: number of end-of-clause pauses ($p = .027$) and lexical density ($p = .031$). Higher proficiency learners did not pause as frequently as their intermediate counterparts did between clauses, but there's no difference between the two proficiency levels in terms of the frequency of mid-clause pauses. Mid-clause pauses were shown to be a trait of L2 speaking (Skehan, 2009), so both high and intermediate proficiency learners in study remained by and large L2 speakers whose oral performance was not much native-like, as far as fluency is concerned. However, the higher proficiency level did appear to lessen the hesitations between clauses. This was probably because a more automatized linguistic system can assemble information in a more coherent manner, making it less likely that the utterances will be disparate or loosely connected to each other.

Having a higher proficiency also resulted in higher lexical density, but the effect size was small (Cohen's d = .39). Higher lexical density is an indicator of more information packed into the same length of speech or text (Halliday, 1993). This result could therefore be explained by the following two possibilities. First of all, a higher proficiency is associated with a bigger vocabulary size. Though not reaching significance, the means of lexical diversity and lexical sophistication all showed a trend that the more advanced learners were in possession of both more different words and more rare words. Since there are only a fixed number function words, a larger vocabulary means more content words, which can contribute to higher lexical density. Second, as mentioned in the last paragraph, participants at a higher level might boast better ability for a more organized speech. As a result, they may get to the point more directly without beating around the bush, hence reducing the use of function words.

However, higher proficiency seemed in most cases not connected to fluency, lexis, and 'nouny' language use (or noun phrase complexity). Past research has showed that fluency and complexity were more easily affected by task-external influences (e.g., planning time), and fluency and complexity were the two places in this study that proficiency had no effect or only a weak effect on. Therefore, with this study and the past literature taken together, a preliminary conclusion is that learner proficiency and task stand in competition. Task(-external) influences would be greater when performance areas are less inherently reflections of proficiency, and visa versa. That said, some research (Skehan and Foster, 1997b, in press; Lee, forthicoming) has begun to show that it is possible to employ pedagogical means such as post-task activities to break this task-proficiency competition.

The above discussion dealt with the main effects of proficiency in different performance areas. Now we will turn to the interaction effects concerning proficiency and see if task (familiarity) types and strategic planning vary according to proficiency levels. Proficiency seems to be a mediating variable in only three places, namely speech rate, clauses per AS unit (complexity) and words per AS unit (complexity), out of all 26 measures employed.

First of all, the three-way interaction in Table 5.13 and figure 5.7 showed that, though in general the differences between familiar and unfamiliar topics in speech rate could be reduced when planning was allowed, such a change was much more significantly achieved within the high proficiency learners. The intermediate learners did benefit from planning to improve their speech rate in both familiar and unfamiliar topics, but there was still a significant difference between them. Learners of high proficiency could speak almost equally fast in either topic after planning. The result indicates a "Matthew effect"[8] in which the higher proficiency participants seemed more able to make the most out of the opportunity to plan and eliminate the adverse conditions induced by their unfamiliarity with the topics. One the other hand, the 'Matthew effect' was not fully observed here because the intermediate learners did not become worse in their speech after planning. They in fact significantly improved, with a smaller margin than their high proficiency counterparts though. The fact that the unfamiliar topics still imposed on them some hindrance is probably a result of their less proceduralized L2 knowledge. They had far more areas to attend to even after planning while speech rate was not always on the top of their performing agenda.

Except for the afore-mentioned familiarity × planning × proficiency 3-way interaction, it is interesting that the general pattern of performance by the two proficiency levels did not seem to be quantitatively differently in the two familiarity task types. It appears that topic familiarity as a kind of task-internal readiness provides a fairly predictable task characteristic for all learners, and needs

[8] The **Matthew effect** (or "accumulated advantage") in sociology is the phenomenon where "the rich get richer and the poor get poorer". Those who possess power and economic or social capital can leverage those resources to gain more power or capital. The term was first coined by sociologist Robert K. Merton in 1968 and takes its name from a line in the biblical Gospel of Matthew: "For to all those who have, more will be given, and they will have an abundance; but from those who have nothing, even what they have will be taken away." – Matthew 25:29, New Revised Standard Version (Wikipedia, June 7, 2010.)

considerations in task design and task-based test fairness.

The other two interactions between proficiency and planning both concern complexity. Wigglesworth (1997) found that the opportunity to plan only allowed learners of higher proficiency, but not those at the lower level, to produce more complex language. Similiarly, Kawauchi's (2005) high proficiency participants benefited most in the case of complexity (and fluency), with the lower proficiency participants gaining less (but they gained the most in accuracy), and the most advanced learners benefiting the least. On the contrary (at first sight), a general pattern from the interactions between planning and proficiency in both 'clauses per AS unit' and 'words per AS unit' in this study is that the intermediate learners were much better than their high proficiency counterparts in making the most out of planning time to achieve higher complexity. For the AS length measure, the difference between high and intermediate participants was narrowed to almost non-existence after planning. More significantly, in terms of the conventional 'clauses per AS unit' measure, the intermediate planners even slightly surpassed the high planners, though the high non-planners were much better than the intermediate non-planners.

Kawauchi (2005) found that learners at a lower level gained the most in accuracy after planning, but Wigglesworth (1997) and Ortega (1999) claimed that planning did help learners at an advanced level for better accuracy in performance. Evidence from this study does not support either side of the inconsistency. No matter whether given planning time or not, the higher learners were always better than the intermediate ones (c.f. the main effects of proficiency above).

Some inconsistency between the present study and the literature in terms of the effects of planning on complexity and accuracy on different proficiency levels may probably be attributed to the operationalization of the independent variable 'proficiency' *per se* (this is not to deny the existence of other possibilities though). It is an age-old problem to equate different proficiency tests in a reliably comparable manner. If the participants of intermediate proficiency in this study are at a level similar to the 'high' participants in Kawauchi (2005) and Wiggleswoth (1997) (and if the 'high' here is equal to the 'advanced' in Kawauchi), then, instead of contradicting, this study could in fact support Kawauchi's results in complexity. That said, such a claim remains a speculation before a commonly acceptable way of equating different proficiency measures is available.

Hypotheses 12 and 13 about the Familiarity × Planning × Proficiency intereaction also revieved general support from the results. Basically, such a three-way interaction did not occur because each of the main effects has served as a stronge driving force in certain performance areas. However, the three-way interaction did occur to one dependeent variable, namely 'words per minute' (speech rate), which shows that students of higher proficiency made better use of planning time to improve their speech rate on the unfamiliar topic than the intermediate students whose speech rate on the unfamiliar topic was still significantly lower than the familiar one even if given planning time. This suggests, once again, another case of the 'Matthew effect' where a learner equipped with better linguistc competence can use the available cognitive resources (more attentional space allowed when given planning time) more efficiently.

6.5 Further discussion: towards a complete picture

The above sections have talked about the three independent variables in a compartmentalized manner. This section, however, will bring together all three independent variables, especially topic familiarity and planning, in the hope of builiding up a clearer picture of task-internal and task-external readiness in L2 task performance. The discussion here will be carried out through a series of fundamental questions that may be essential in understanding the general framework of learner readiness.

6.5.1 Topic familiarity VS Strategic planning, which is more powerful?

A question asked at the beginning of this chapter concerning the general framework of learner readiness was: task-internal readiness (e.g., topic familiarity) or task-external readiness (e.g., planning): which is a stronger driving force for the improvement of task performance? It was discussed in Section 7.4 that proficiency in most cases does not interfere with the effects of topic familiarity and planning.

Table 6.3 Effect sizes produced by topic familiarity and strategic planning

	Topic familiarity	Strategic planning
Pruned total words	.61	.44
Breakdown fluency	.17 – .38	.32 – .71
Repair fluency	.31 – .40	.43 – 1.02
F-score ("Nouniness")	.49	.48
Complexity	*NS*	.39 – .81
Lexical density	*NS*	.43
Accuracy	.22 – .38	*NS*
Lexical diversity	.29	*NS*
Lexical sophistication	.41	*NS*

Then, though proficiency may not be a highly interesting variable, the situations have made a clearer comparison between topic familiarity and strategic planning possible.

Table 6.3 above sums up the effect sizes produced by the two independent variables. The reason for choosing effect sizes is obvious: effect sizes highlight the magnitude of the independent variables' effects. Also, its existence *per se* indicates that there is a significant effect. Bearing this in mind, we can figure out that familiarity and planning displayed very similar patterns in meaning expression aspects: total words, breakdown fluency, repair fluency and also F-score ('nouniness'); but they differed in the formal or organizational aspects (accuracy and complexity) as well as lexical aspects. Another interesting feature is that topic familiarity showed small effects in most cases, whereas planning generally produced much higher values of Cohen's *d*. Take breakdown fluency as an example, effect sizes for planning were almost always nearly twice as high as those for topic familiarity. Furthermore, in the formal features of L2 speaking, the effect sizes in complexity produced by planning were also much higher than those in accuracy by topic familiarity. Therefore, an answer to the question about which is a stronger variable seems to be emerging: planning, or task-external manipulation, appears to be a more powerful influence on task performance than topic familiarity, a task-internal variable. Hypothesis 7 that being familiar with a topic and the opportunity to plan will have equal strength in influencing fluency was rejected.

Some recap on the general framework in section 6.1 would probably help to explain this pattern. Task-internal readiness (including topic familiarity, schematic familiarity, and task familiarity) is some sort of implicit or unconscious preparation that a learner brings to a task which will function both before and during the actual performance. An important characteristic of task-internal readiness is that learners are not necessarily aware of what privilege they enjoy. In contrast, task-external readiness (i.e. rehearsal, strategic planning, and online planning) provides an explicit and announced push for learners to be embraced for the subsequent tasks. It would be therefore fair to say that task-external readiness constitutes a greater extent in preparation and thus becomes more powerful in areas that it has influence on than task-internal readiness.

6.5.2 Is task-internal readiness less useful than task-external readiness?

Though strategic planning has greater effects on fluency and complexity than topic familiarity, topic familiarity nevertheless was more able to push learners in accuracy and lexical aspects. Therefore, task-internal readiness functions in different areas from task-external readiness. What attracts our attention is that topic familiarity seemed to enable participants to mildly but still significantly strike a balance between meaning and form, which might signal an integration of their linguistic knowledge into genuine performance. Bygate and Samuda (2005) pointed out that 'a common learning and teaching problem is to get learners to integrate knowledge that is available to them into their active language use (p37)'. In this sense, providing learners with familiar topics to practise may better encourage them to conform to this pedagogical end.

Strategic planning promotes pre-task readiness while on-line planning results in real time preparation. Though much less powerful in comparison to each of these two task-external means alone, task-internal readiness appears to consist of the features of pre- and during-task readiness as it is inherent within each learner and could take effects both prior to and during a task. As discussed earlier, the on-line readiness nature of topic familiarity may probably contribute to the better accuracy performance. Therefore the integration of linguistic knowledge into communicative use could be one important area in exploring task-internal readiness in future.

6.5.3 Can task-internal and task-external readiness compensate for each other?

Intuitively we might expect a compensation effect for task-internal and task-external readiness. For example, one can think of providing a familiar topic for non-planners in the hope that they can speak as fluently as those planners performing an unfamiliar topic. Data from the present study did reveal such a compensation effect in breakdown fluency (thus lending support to Hypothesis 8), but the conclusion is that planning could compensate for the unfamiliar topics much better than familiar topics could do for the unplanned conditions. In five breakdown measures, planners reached almost the same fluency level in both familiar and unfamiliar topics. The adverse condition in fluency induced by their lack of domain knowledge was clearly removed when planning time was offered. However, the significant difference between planners and non-planners continued to exist even after familiar topics were performed. This result echoes the above discussion that task-external manipulation is a stronger driving force for many areas, especially fluency.

However, the compensation effects happened almost all in fluency measures only, Hypthosis 9 about the interactions in accuracy and compelxity did not seem to be confirmed.

6.5.4 What are the pedagogical implications from the results?

The pedagogical implication regarding task-external readiness (e.g., strategic planning) has been researched in many studies (see Ellis, 2005 and in press, for a detailed discussion), but the benefit of using task-internal readiness has rarely been examined in the literature. Evidence from this study, however, showed that task-internal readiness should not be ignored in language education.

First of all, previous research has shown that *receptive* language use, namely reading comprehension (Shimioda, 1993; Chang, 2006; Lee, 2007; Leeser, 2007; Barry and Lazarte, 1995; Bügel and Buunk, 1996; Chen and Donin, 1997; Johnson, 1982; Lee, 1986) and listening comprehension (Markham and Latham, 1987; Long, 1990; Chiang and Dunkel, 1992; Schimidt-Rinehart, 1994; Leeser, 2004), is greatly influenced by one's familiarity with background knowledge. This study further provides evidence for the the effects of familiarity in L2 speech production, as *productive* language use. Familiarity may therefore become an inevitable issue in test fairness. It is highly possible that one performs well not because s/he is in fact more proficient but simply because s/he is more familiar with the topic. Match and mismatch between test content and learner's background have to be taken into serious consideration in either language comprehension or production tests.

Secondly, one of the important issues in task-based language instruction is to encourage learners to participate actively in various task activities. This study shows that, providing learners more familiar topics will reduce learner anxiety and elevate their willingness to communicate, as evidenced by the significantly longer account on familiar topics. A longer speech produced by an L2 learner is an indication of his/her willingness and readiness to communicate on the one hand. On the other hand, this certainly helps to enhance learner confidence, which may work especially for low to intermediate learners.

Thirdly, strategic planning was shown to be able to help learners with more fluent and more complex language. Therefore, it would appear to be a good idea to allow learners some time pior to the actual performance. Planning would encourage learners to embark on more elaborated language, venturing more complex structures through which they could experiment the newly aquired linguistic knowledge. Planning also serves to narrow the gaps between high and low proficiency, and between familiar and unfamiliar tasks, in terms of fluency and complexity. In classrooms, then, teachers may take advantage of planning when learners are facing adverse situations (such as low profiency and unfamiliar topics).

Fourthly, the results suggest that, learners should be provided with familiar topics in tasks if lexis and accuracy are the concerns. It is not uncommon that students learn a lot of new words but when it comes to actual speaking, they tend to reply on a very limited set of vocabulary with which they are most familiar. It would then seem that this dead knot could be tackled by creating familiar tasks though which learners have the opportunity of drawing on a larger, and more diverse, vocabulary from their mental lexicon. Given the nature of familiar as mini online planning, familiarity helps learners also with accuracy as they have more resources to attend to form. As mentioned above, this may increase their confidence and reduce the feeling of frustration.

Fifthy, this study also has implications for task sequencing. We have seen the separate benefits for pedagogy from each individual variable, but it is far more important to examin how these bits and pieces are organized to form a coherent and organic whole. It is certainly too early to make any claims on the "whole picture" based on the three variables in this study alone. Nonetheless, this study indicates that at the pre-task stage planning is a useful tool, whilst at the during-task stages familiarity may help. Then, beginners should receive the most familiar topics and planning time in order that they could be fully supported in tasks. As their language ability develops, either familiarity or planning could be removed from the favourable conditions so that they would face greater (but appropriate) chanllenges and be motivated to proceed further.

Sixthly, the results concerning profiency point out that, though task manipulations can improve performance in terms of accuracy or complexity, one's own proficiency is the best indicator or predictor of one's underlying syntactic abilities. In language testing, therefore, we can still have confidence in the performance as one's profiency (especially in syntactic areas), provided that task conditions and task characteristics are held constant. In addition, because learners of lower proficiency can approach the level of higher proficiency learners in some performance areas when given appropriate task support (e.g., planning helps in syntactic complexity and topic familiarity helps in lexis), we should fully explore the potential of tasks in helping the underachievers in future.

Last but not the least, the present study supports content-based instruction (Mohan, 1986) in language teaching. Topic familiarity proved to be a positive influence on fluency, accuracy and lexis, with indications that it helps to push learners to a more integrative approach to language learning. Compared to 'pure' or intensive language teaching, language seems more effectively taught when the domain knowledge (not linguistics knowledge) is imparted to learners in their L2, leading to their genuine need to solve real world problems. In a language classroom where general knowledge is not the focus, language can still be taught using tasks involving connections to real life so that tasks become the medium between classroom and actual society (c.f. Skehan 1998 for the definition of a task).

6.5.5 What new performance constructs can be abstracted from the present data set?

In connection to section 5.10 about the results of the factor analyses, a number of theoretical possibilities in performance constructs seem to be emerging. First, a new construct on fluency that is connected to pauses at the end of clause appears to have emerged. Skehan (2009) pointed out that the position of pauses is an important trait to distinguish native from non-native speakers because native speakers tend to pause at the end of a clause to facilitate listening comprehension, but non-native speakers pause where they have to, in both mid-clause and end-of-clause positions. For a listener, the end of a clause is a more natural place to pause and sometimes such pauses even turn out to be unnoticeable. Identifying such a fluency construct that does not cluster with either breakdown fluency or repair fluency seems to reveal two implications. On the one hand, this construct was abstracted from both tasks and may indicates a more native-like performance in fluency because typical L2 learners pause everywhere (Skehan, 2009) and do not necessarily differentiate between mid-clause and end-of-clause pauses. If end-of-clause pauses achieve an outstanding position, they may then become the evidence of approaching native-like performance of fluency. On the other hand, this result shows that fluency is a multi-dimensional construct that might consist of more aspects than we previously thought (previous research generally identified speech, breakdown fluency, and repair fluency. See Ellis, 2005). This would hopefully invoke future research.

Secondly, the construct of repair fluency seems rather unstable. In this study, though planning reduced the number of false starts, reformulation and repetitions, it also pushed learners to more replacements. This may echo what Kormos (2006) suggested that repairs in L2 speech are 'good indicators' of the encoding processes which have not yet become fully automatized. On the one hand, repairs may show hesitations in speech which results in dysfluency from a temporal perspective; on the other hand, repairs could also indicate the tendency for better language, such revising errors or replacing a word with a better choice. In this sense, the less clear repair fluency construct could derive from its multi-faceted connotations.

Thirdly, this study appears to have identified a 'noun phrase complexity' construct involving four measures: words per clause, F-score, DB-score and lexical sophistication. The close link between clause length and 'nouny' language seems plausible in that the size of the noun phrase could contribute more to the clause length than verb phrases whose structure is relatively less expandable. The DB-score is negatively associated with the rest because the lower the score is, the more informational the language will become (Biber et al, 1998). Given the fact that information is largely carried on by noun phrases instead of verb phrases, the DB-score found its rightful position in noun phrase complexity as well.

The interesting measure here is lexical sophistication, as indexed by the value of *Lambda*, which denotes the depth of lexical use, or the extent to which more rare words are utilized. Lexical sophistication looks, at first sight, irrelevant to the size of a noun phrase or clause length. Rather, it seems to indicate the quality of the noun phrase. The positive values in the loading in the factor analysis indicate that learners (at least at this relatively higher proficiency level) expand the noun phrase structure through the use of more low frequency words. This could also suggest that words of lower frequency drives for a more complex noun phrase structure in the discourse. Skehan (2009) argued that more demanding lexis leads to more complex syntax among native speakers, but can disrupt syntactic planning among non-native speakers, which results in a trade-off between syntactic complexity and lexical complexity.

Compared to the non-native speakers in the six studies analyzed in Skehan (2009) who were mostly beginners to lower intermediate learners, the more advanced learners in this study seemed to show a certain degree of native-like tendency to have noun phrase complexity and lexical complexity come in tandem. That said, syntactic complexity, as measured by clauses per AS unit, did not show much connection with lexical sophistication. This is consistent with Skehan (2009) and indicates that participants in the present study were still non-native speakers in this area. A preliminary conclusion is available at this point: higher proficiency appears to play a role in easing trade-off effects prevalent in L2 learners. Though we may agree what there is qualitative distinction between native and non-native speakers, there seems to be a continuum towards native-like proficiency along which non-nativeness decreases.

Last but not least, the measure of 'words per AS unit' seems to be divided between its intended construct, syntactic complexity, and the unexpected category, breakdown fluency. It is natural to see the conventional complexity measure 'clauses per AS unit' closely linked to this length of AS unit measure because the higher subordination ratio there is, the longer the sentence will usually be. The curious point is how the length of an AS unit also comes to share a moderate loading with other breakdown fluency measures. Breakdown fluency seems to be a Formulator function in Levelt's (1989) model where speakers are engaged with a second-by-second decision. The pauses will interrupt the ongoing speech which leads to more breakdowns. In this study, more fluent language seems to be able to increase the length of an AS unit. This is intriguing but could be explained by an "inertia" theory borrowed from physics. If a L2 speaker is fluent, s/he could go on with the sentence to where s/he has to stop. On the contrary, when the speaker encounters breakdowns, it is more likely that s/he would have to start another utterance and leave behind a short AS unit, than when s/he is fluent and has the whole utterance completed.

Chapter 7 Conclusion, Limitations, and Future Directions

The present chapter revists the research questions, summarizes key findings, and reflects on the limitations of the study. Based on the findings and the limiations, directions are suggested for future research. A final remark at the end concludes the entire research.

7.1 Summary of the research

The past twenty years has seen rapid development in task-based language learning research. Whereas some generalizations can be made on the basis of empirical data available (see, for example, Ellis, in press, for a review of the findings on planning), inconsistency exists in other areas. Moreover, there is still a vast virgin land left untouched. The purpose of this study then was two-folded: to replicate the previous research in strategic planning, and to explore one new area (topic familiarity) in L2 speaking, with their individual and interaction effects on different proficiency levels, which is also a less researched variable in TBLT. Such a combination yielded a range of interesting findings, among which the most important aspects are highlighted as follows:

1) The concept of planning can be extended to task-readiness which involves two macro dimensions: task-internal readiness and task external readiness, each

with three micro dimensions. Task-internal readiness further subsumes topic familiarity, schematic familiarity, and task type familiarity, while task-external readiness includes rehearsal, strategic planning, and online planning. This proposal of the general framework of task-readiness can potentially serve as a theoretical platform to unify and synthesize research in various types of planning, familiarity, and even other kinds of preparatory activities for a task.

2) Both planning and topic familiarity raise fluency, indicating that participants with task-readiness prioritize meaning expression. When planning or topic familiarity is present, proficiency appears to be largely overridden in its effect on fluency.

3) Planning produces bigger effect sizes than topic familiarity in fluency. Planning is also able to greatly reduce the gap between familiar and unfamiliar topics in fluency. This leads us to the conclusion that task-external readiness is more powerful than task-internal readiness in improving meaning-oriented performance.

4) Planning raises syntactic complexity, while topic familiarity increased accuracy. It would then appear that task-internal readiness encourages learners to a conservative stance (thus higher accuracy), but task-external readiness pushes learners to task risks (hence higher complexity). Interestingly, higher proficiency produces much higher accuracy and moderately higher complexity, confirming a close relation between syntactic performance and linguistic competence.

5) With the above points taken together, an intriguing pattern emerges – task

influence and proficiency influence does not always complement each other. The proficiency-oriented variables (e.g., accuracy) are affected more by proficiency levels and less by task manipulations, whereas task-oriented variables (e.g., fluency) function just on the opposite. There are also intermediate variables, such as complexity.

6) Topic familiarity drives learners to produce higher lexical diversity and higher lexical sophistication, while planning and proficiency have effects on lexical density. This indicates that one's prior knowledge with a certain subject is associated with a bigger pool of *productive* lexis (so that they recycle less) which also include more rare words. Planning time enables learners to retrieve more information from the memory store and results in a higher ratio of content words.

7) Some insights are also available for the performance measures. First of all, this study has identified an end-of-clause fluency construct which distinguishes itself from the recognized breakdown fluency and repair fluency. Secondly, there seems to be a noun phrase complexity construct which is different from syntactic complexity. Noun phrase complexity is established on the basis of clause length and noun phrase use. All this may have implications for future performance measurement.

7.2 Limitations

Though every effort has been made to improve the quality of this research, several limitations have to be acknowledged. The first limitation lies in the unequal proficiency levels between the medicine majors and the computer majors, as discussed in the methodology section. Even if this problem was solved through the use of statistical procedures (MANCOVA), the unequal levels make the comparisons between different groups less straightforward to interpret. Greater control of the variables would have been desirable.

Secondly, this study was based on a purely quantitative paradigm. Therefore some interesting individual differences might have been covered up by examining the group means alone. Also, some information about the actual psycholinguistic processes during task performance was not gathered.

Thirdly, though the C-test has been proved in this and many other studies as an effective measure to differentiate learners in terms of their proficiency levels, the equation between the C-test and other established proficiency tests is less clear. It worked well for within-study proficiency distinction for the purpose of grouping, but it is then difficult to conduct cross-study comparisons as far as a same proficiency level is concerned. Dornyei and Katona (1992) was a good start to solve this problem as they compared the C-test with a conventional cloze test and a TOEIC test. However, their focus was to find correlations between the tests (for concurrent validity) and did not take the step necessary to equate the tests.

Finally, this study approached task performance from an analytic point of view by examining a host of precise and detailed measures. Though this method helps with our understading of various specific areas in task performance, information about a wholistic picture, i.e., how comprehensible the performances were and how well each participant actually completed the tasks, was far less touched upon. It is possible that a participant speaks fluently and accurately but s/he in fact strays away from the topic.

7.3 Future directions

Future research shall focus on two main directions: the rectification of the limitations indentified in the last section, and the extension of the present research based on the trends revealed by the findings. For the first direction, future studies should recruit more participants from whom various comparable groups (in terms of proficiency) can be formed. Also, post-task interviews and during-task think-aloud protocols may be employed to gain insights into the qualitative information about on-going task performance. As for the comparison of the C-test to other more widely used proficiency tests, the best hope certainly lies in testing experts carrying out serious research in equating C-test scores with other measures (e.g., IELTS and TOEFL). For a task researcher, a concurrent validation process is strongly recommended in order to see the relationship between C-test and other tests.

The second main direction is concerned with extending the present studies in greater width and depth. A general framework of task-readiness (see section 6.1) was proposed based on one task-external readiness variable (strategic planning) and one task-internal readiness variable (topic familiarity) in this study, with other dimensions added according to the literature. Then an obvious direction for greater width is to have a study that takes in all three types of task-internal readiness (topic familiarity, schematic familiarity, and task type familiarity) and all three types of task-external readiness (rehearsal, strategic planning, and online planning). Then a clearer conclusion can be drawn for the effects of all six types of task-readiness and their pedagogical implications.

Another direction for follow-up research is to conduct studies at greater depth. For example, this study involved mainly two proficiency levels, namely intermediate and high. Future studies can include an additional group at a low proficiency level and perhaps clearer proficiency effects can be borne out. Another possibility is to explore ways of compensating for adverse conditions induced by learners' unfamiliarity with the topics. In the research context and the classroom context alike, L2 learners are frequently asked to do tasks that they have never encountered, so unfamiliarity is perhaps the norm with tasks instead of familiarity. As we have seen in the results, planning is able to narrow the gap between familiarity types as far as fluency is concerned. An emerging question is then: what other means could affect other performance areas (like accuracy and complexity)? One suggestion is to provide learners with input such as reading materials so as to help them get familiar with topics. This may then enable the reading–speaking connection to be explored in an interesting way.

In addition to the above major directions, experienced oral test rators must be invited in future studies to judge the comprehensibility of the speaking as well as the extent of task completion. Such an arrangement adds valuable information about the overview of the task performance to the examination of each specific perfomace area by the analytic measures.

7.4. Conclusion

As has been said in the Introduction chapter, it is not too easy to draw indubitable conclusion in social sciences, and I admit that this assertion also applies to the present study. That being said, this study was conducted in the context of TBLT research and established very close connections to prior studies, thus enabling cross-study comparisons. It is then the hope that this research will be a link between the literatures on planning and the future studies on the extended concept of planning, task-readiness, to explore task-based language learning from an even wider perspective.

References

Arras, U., Eckes, T. & Grotjahn, R. (2002). C-Tests in the Rahmen des 'Tests Deutsch als Fremdsprache' (TestDaF) : Erste Forschungsergebnisse. In Grotjahn R. (Ed.). *Der C-Test. Theoretische Grundlangen und praktische anwendungen* (vol. 4) (pp. 175-210). Bochum: AKS-Verlag.

Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.

Anderson, R. C. & Pearson, P. D. (1984). A schema theoretic view of basic process in reading comprehension. In P. D. Pearson (Ed.), *Handbook of Reading Research* (pp. 255-291). New York: Longman.

Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Baddeley, A. (1997). *Human Memory: Theory and Practice*. New York: Psychology Press.

Baddeley, A. & Hitch, G. (1974). Working momory. In G. Bower (Ed.). *The Psychology of Learning and Motivation*. (Vol. 8) (pp.47-89). New York: Academic Press.

Banks, J. (2004). The Impact of event familiarity on the complexity and coherence of children narratives of positive events. Unpublished MSc Thesis. North Carolina State University.

Barry, S. & Lazarte, A. (1995). Embedded clause effects on recall: Does high prior knowledge of content domain overcome syntactic complexity in students of Spanish? *Modern Language Journal*, *79*, 491–504.

Bartlett, F. C. (1932/1977). *Remembering-A Study in Experimental and Social Psychology.* Cambridge: Cambridge University Press.

Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus Linguistics.* Cambridge, England: Cambridge University Press.

Bock, K. & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, 44 (2), 123-147.

Brown, G. & Yule, G. (1983). *Discourse Analysis.* Cambridge: Cambridge University Press.

Bruton, A. (2002). From tasking purposes to purposing tasks. *ELT journal*, 56 (3), 1-30.

Bügel, K. & Buunk, B. (1996). Sex differences in foreign language text comprehension: The role of interests and prior knowledge. *Modern Language Journal*, 80, 15–31.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and Change in*

Language Teaching (pp. 136-146). Oxford: Heinemann.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds). *Researching Pedagogicl Tasks: Second Language Learning, Teaching and Testing* (pp.23-48). Harlow, England: Longman.

Bygate, M. & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 37-76). Amsterdam: Johan Benjamins.

Canale, M. (1983). On some dimensions of language proficiency. In J. W. Oller (Ed.). *Issues in Language Testing Research* (pp. 333-342). Rowley, Mass.: Newbury House.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.

Carrell, P. L. (1983). Three components of background knowledge in reading comprehension. *Language Learning*, 33, 183-207.

Carrell, P. & Eisterhold, J. (1983) Schema theory and ESL reading pedagogy. *TESOL Quarterly*, 17, 553–573.

Carroll, D. W. (1999). *Psychology of Language* (3rd Ed.). Pacific Grove, CA: Brooks/Cole.

Chang, C. (2006). Effects of topic familiarity and linguistic difficulty on the reading strategies and mental representations of non-native readers of Chinese. *Journal of Language and Learning*, 4, 172-198.

Chang, Y. F. (1999). Discourse topics and interlanguage variation. In P. Robinson (Ed.). *Representation and Process: Proceedings of the 3rd Pacific Second Language Research Forum* (Vol.1) (pp. 235-241). Tokyo: PacSLRF.

Chen, Q. & Donin, J. (1997). Discourse processing of first and second language biology texts: Effects of language proficiency and domain-specific knowledge. *Modern Language Journal*, *81*, 209–227.

Chern, C. (1993). Chinese students' word-solving strategies in reading in English. In T. Huckin, M. Haynes, & J. Coady (Eds.). *Second language reading and vocabulary learning* (pp. 67–82). Westport, CT: Ablex.

Chiang, C. S. & Dunkel, P. (1992). The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL Quarterly*, 26, 345-373.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, Mass.: MIT Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Cohen, A. D., Segal, M. & Bar-Siman-Tov, R. (1984). The C-Test in Hebrew. *Language Testing*, 1 (2), 221-225.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11, 367-383.*

Dai, B. (2009). Re-examining individual differences in working momory, learner awareness of L2 forms and L2 development through recasts on task-based interaction. Paper presented at the 3rd Biennial International Conference on Task-Based Language Teaching, Lancaster University, U.K.

Daller, H. & Phelan, D. (2006). The C-test and TOEIC as measurements of students' progress in intensive short courses in EFL. In Grotjahn, R. (Ed.). *Der C-Test: Theorie, Empirie, Anwendungen. (The C-Test, Theory, Empirical Research, Applications)* (pp. 101-119). Frankfurt am Main: Lang.

Daller, H. & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students, in Daller, H., Milton, J. and Treffers-Daller, J. (Eds). *Modelling and Assessing Vocabulary Knowledge* (pp. 150-164). Cambridge: Cambridge University Press.

De Bot, K. (1992). A bilingual production model: Levelt's 'Speaking Model' adapted. *Applied Linguistics*, 13, 1-24.

De Bot, K. & Schreuder, R. (1993). Word production and the bilingual lexicon. In R.Schreuder & B. Weltens (Eds.). *The Bilingual Lexicon.* (pp. 191-214). Amsterdam: John Benjamins.

Dornyei, Z. & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, 9, 187-206.

Douglas, D. (2000). *Assessing Languages for Specific Purposes.* Cambridge: Cambridge University Press.

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. (2005). Planning and task-based performance: theory and research. In R.

Ellis (Ed.). *Planning and task performance in a second language* (pp. 3-36). Amsterdam: Johan Benjamins.

Ellis, R. (in press). The Differential Effects of Three Types of Task Planning on the Fluency, Complexity and Accuracy in L2 Oral Production. *Applied Linguistics.*

Elllis, R. & Barkhuizen, G. (2005). *Analysing Learner Language.* Oxford: Oxford University Press.

Ellis, R. & Yuan, F. Y. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.). *Planning and task performance in a second language* (pp. 167-192). Amsterdam: John Benjamins.

Eysenek, M. W. & Keane, M. T. (1990). *Cognitive Psychology: A Student' s Handbook.* Hillsdale, NJ : Lawrence Erlbaum Associates.

Færch, C. & Kasper, G. (1983). *Strategies in Interlanguage Communication.* London: Longman.

Foster, P. (2001). Rules and routines: a consideration of their role in task-based language production of native and non-native speakers. In Bygate, M., Skehan, P., & Swain, M. (Eds.), *Researching Pedagogic Tasks: Teaching, Learning and Testing* (pp. 75-97). Longman, London.

Freed, B. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbach (Ed.), *Perpsectives on Fluency* (pp. 243-265). Ann Arbor MI: the University of Michigan Press.

Gardner, R. C. (2001). *Psychological Statistics Using SPSS for Windows*. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Garman, M. A. (1990). *Psycholinguistics*. Cambridge: Cambridge University Press.

Garrett, M. F. (1981). Levels of processing in sentence production. In B. Butterworth (Ed.). *Language Production* (Vol.1): *Speech and Talk* (pp. 177-200). London: Academic Press.

Good D. A. & Butterworth, B. (1980). Hesitancy as a conversational resource: some methodological implications. In H. Dechert and M. Raupach (Eds.). *Temporal Variables in Speech Production* (145-152). The Hague: Mouton.

Grotjahn, R. (1995). Der C-Test: State of the art. *Zeitschrift fur Fremdsprachenforschung*, 6 (2), 37-60.

Halliday, M. A. K. (1970). Functional diversity in language as seen from a consideration of modality and mood in English. *Foundations of Language*, 6, 322-361.

Halliday, M. A. K. (1976). The form of a functional grammar. In Gtinter Kress (Ed.). *Halliday: System and function in language* (pp. 7-25). Oxford: Oxford University Press.

Halliday, M. A. K. (1994). *An introduction to Functional Grammar* (2nd Ed.). London: Arnold.

Halliday, M. A. K. & Martine, J. R. (1993). *Writing science: Literacy and disursive power*. Pittsburgh: University of Pittsburgh Press.

Hammadou, J. (1991). Interrelationships among prior knowledge, inference, and language proficiency in foreign language reading. *The Modern Language Journal*, 75, 27-39.

Heylighen, F. & Dewaele, J. (1999). Formality of language: Definition, measurement and behavioral determinants. Internal report, Center "Leo Apostel", Free University of Brussels.

Hudson, T. (1982). The effects of induced schemata on the "short-circuit" in L2 reading: non-decoding factors in L2 reading performance. *Language Learning*, 32(1), 1-31.

Hymes, D. (1972). On communicative competence. In J. Pride and J. Holmes (Eds.). *Sociolinguistics* (pp. 53-73). Harmondsworth: Penguin.

Johnson, P. (1982). Effects on reading comprehension of language complexity and cultural background of text. *TESOL Quarterly*, 16, 169–181.

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36.

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate proficiency. In R. Ellis. (Ed.). *Planning and task performance in a second language* (pp. 143-164). Amsterdam: John Benjamins.

Kintsch, W. (1988). The role of knowledge of discourse comprehension: A construction-integration model. *Psychological Review*, 92, 163-182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Klein-Braley, C. (1985). A cloze-up on the C-Test: A Study in the Construct Validation of Authentic Tests. *Language Testing*, 2 (1), 76-104.

Klein-Braley, C. (1997). C-test in the context of reduced redundancy testing: an appraisal. *Language Testing*, 14 (1), 47-84.

Klein-Braley, C. & Raatz U. (1984). A survey of research on the C-Test. *Language Testing*, 1 (2), 134-146.

Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Lee, J. F. (1986). Background knowledge and L2 reading. *Modern Language Journal*, 70, 350–354.

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students reading comprehension and learning of passive form. *Language learning*, 57, 87-118.

Lee, J., & Schallert, D. L. 1997. The Relative Contribution of L2 Language Proficiency and Ll Reading Ability to L2 Reading Perfirmance: A Test of the Threshold Hypothesis in an EFL Context. *TESOL Quarterly*, 31(4), 713-739.

Leeser, M. J. (2003). *Second language comprehension and processing grammatical form: The effects of topic familiarity, mode, and pausing*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Leeser, M. J. (2004). The effects of topic familiarity, mode, and pausing on second language learnres' comprehension and focus on form. *Studies in*

Second Language Acquisition, 26, 587-615.

Leeser, M. J. (2007). Learner-Based Factors in L2 Reading Comprehension and Processing Grammatical Form: Topic Familiarity and Working Memory. *Language Learning*, 57 (2), 229 – 270.

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

Levelt, W. J. M. (1993). *Speaking: From intention to articulation.* Cambridge MA: The MIT Press.

Levelt, W. J. M. (1999a). Models of Word Production. *Trends in Cognitive Sciences*, 3(6), 223 – 232.

Levelt, W. J. M. (1999b). Language production: A blue print of the speaker. In C. Brown & P. Hagoort (Eds.). *Neurocognition of language* (pp. 83-122). Oxford, England: Oxford University Presss.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (2000). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1 – 75.

Li, E. C., Williams, S. E, & Volpe, A. D. (1995). The effects of topic and listener familiarity on discourse variables in procedural and narrative discourse tasks. *Journal of Communication Disorders*, 28, 39 – 55.

Long, D. R. (1990). What you don't know can't help you: an exploratory study of background knowledge and second language listening comprehension. *Studies in Second Language Acquisition*, 12, 65 – 80.

Lynch, T. & J. Maclean. (2000). Exploring the benefits of task repetition and recyling for classroom language learning. *Language Teaching Research* 4, 221—250.

Lynch, T. & J. Maclean. (2001). Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.). *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 141-162). Harlow, England: Longman.

Malvern, D. & Richards, B. (2002). Investigating accommodation in a language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104.

Markham, P. & Latham, M. (1987). The influence of religious-specific background knowledge on the listening comprehension of adult second-language students. *Language Learning*, 37, 157 – 170.

Matthiessen, C. and Halliday, M. K. (1997). Systemic functional grammar: a first step into the theory. Retrieved from: http://minerva.ling.mq.edu.au/resource/VirtuallLibrary/Publications/sfg_firststep/SFG%20intro%20New.html . Oct. 29, 2009.

Meara, P. and Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16 (3), 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.

Merlo, S., & Mansur, L. L. (2004). Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders*, 37, 489-503.

Meyer, L. S., Gamst, G.., & Guarino, A. J. (2006) *Applied Multivariate Research: Design and Interpretation*. Thousand Oaks: Sage Publications.

Minsky, M. A. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The Psychology of Computer Vision* (pp. 211-277). New York: McGraw-Hill.

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.

Mohan, B. A. (1986). *Language and Content*. Mass.: Addison-Wesley Publishing Company.

Nassaji, H. (2002). Schema theory and knowledge-based processes in second language reading comprehension: A need for alternative perspectives. *Language Learning*, 52, 439–481.

Norris, S. P., & Phillips, L. M. (1987). Explanation of reading comprehension: Schema theory and critical thinking theory. *Teachers College Record*, 89, 281-306.

Ortega, L. (1995). The effects of planning in L2 Spanish narratives. *Research Note* 15. Honolulu H1: University of Hawai'I, Second language Teaching and Curriculum Center.

Ortega L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109 – 148.

Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.). *Planning and Task Performance in a Second Language* (pp. 77-109). Amsterdam: John Benjamins.

Ortega, L., Iwashita, N., Norris, J., & Rabie, S. (in preparation). *A multi-language comparison of syntacticcomplexity measures and their relationships to foreign language proficiency*. Manuscript in preparation.

Pang, F. & Skehan, P. (in preparation). What do learners do when they plan? In P. Skehan (Ed.). *Contextualising Task-based Research: A Processing Perspective.* Amsterdam: John Benjamins (proposed).

Payne, J. S. (2000). A study of the effects of individual differences in working memory capacity and synchronous computer mediated communication in a second language on second language oral proficiency development. Unpublished Ph.D. dissertation. Washington State University. U.S.A. Retrieved from https://research.wsulibs.wsu.edu:8443/jspui/bitstream/2376/37/1/j_payne_050100.pdf , Nov. 5, 2009.

Peretz, A., & Shoham, M. (1990). Testing reading comprehension in LSP: Does topic familiarity affect assessed difficulty and actual performance? *Reading in a Foreign Language*, 7, 447–455.

Phillips, L. M. (1990). *Inference Strategies in Reading Comprehension*. Center for the Study of Reading Technical Report No. 410. Champaign, IL: University of Illonois.

Piaget, J. (1970). *Genetic epistemology.* (E. Duck worth, Trans.). New York: Columbia University Press.

Poulisse, N. (1997). Language production in bilinguals. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum.

Read, J. (2000). *Assessing Vocabulary.* Cambridge: Cambridge University Press.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language learning*, 45, 99 – 140.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions. *Studies in Second Language Acquisition,* 18(1), 27-67.

Robinson, P. (2001a). Task complexity, task difficulty and task production: exploring interactions in a componential framework. *Applied Linguistics*, 22, 27 – 57.

Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influence on SLA. In P. Ronbinson (Ed.). *Cognition and Second Language Instruction.* Cambridge: Cambridge University Presss.

Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan, and M. Swain (Eds). *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 119-140). London, UK: Longman.

Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.). *Planning and Task Performance in a Second Language* (pp. 111-141). Amsterdam: John Benjamins.

Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people.* New York: Cambridge University Press.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Hillsdale, NJ: Erlbaum.

Schmidt-Rinehart, B. C. (1994). The effects of topic familiarity on second language listening comprehension. *The Modern Language Journal*, 78, 179-189.

Sheen, R. (1994). A critical analysis of the advocacy of the task-based syllabus. *TESOL quarterly*, 28 (1), 127-51.

Shimoda, T. A. (1993). The effects of interesting examples and topic familiarity on text comprehension, attention, and reading speed. *Journal of Experimental Education*, 61, 93-103.

Skehan P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics,* 17 (1), 38-62

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, and M. Swain (Eds.). *Researching pedagogic tasks: Second language learning, teaching and testing*. London, UK: Longman.

Skehan, P. (2009). Lexical performance by native and non-native speakers on language learning tasks. In Helmut Daller, David D. Malvern, Paul Meara, James Milton, Brian Richards, Jeanine Treffers-Daller (Eds.). *Vocabulary Studies in First and Second Language Acquisition: The Interface between Theory and Application*. London: Palgrave Macmillan.

Skehan, P. & P. Foster. (1997a). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1 (3), 185-211.

Skehan, P. & Foster, P. (1997b). The influence of planning and post-task activities on accuracy and complexity in task-based learning. *Language Teaching Research* 1/3.

Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions in narrative retellings. *Language Learning*, 49 (1), 93-120.

Skehan, P. & Foster, P. (2005). Strategic and online planning: the influence of surprise information and task time on second language performance. In R. Ellis. (Ed.). *Planning and task performance in a second language* (pp. 193-216). Amsterdam: Johan Benjamins.

Skehan, P. & Foster, P. (2007). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In Van Daele, S., A.F. Housen, F.M. Kuiken, M. Pierrard and I. Vedder (Eds.). *Complexity, Fluency and Accuracy and Fluency in Second Language Use, Learning and Teaching* (pp.207 - 226). Brussels: VWK.

Van Daele, S., A.F. Housen, F.M. Kuiken, M. Pierrard and I. Vedder

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, 27 (3), 363 – 374.

Swan, M. (2005). Legislation by hypothesis: the case of task-based instruction. *Applied linguistics*, 26 (3), 376-401.

Tajima, M. (2003). The effects of planning on oral performance of Japanese as a foreign language. Unpublished dissertation. Perdue University.

Tavakoli, P. & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis. (Ed). *Planning and task performance in a second language* (pp. 239-273). Amsterdam: Johan Benjamins.

Thalheimer, W., & Cook, S. (2002). *How to calculate effect sizes from published research articles: A simplified methodology.* Retrieved April 18, 2009 from http://work-learning.com/effect_sizes.htm.

Trammell, R. L. 1991. Can the ACTFL proficiency guidelines be used for reading for-research courses? *ADFL Bulletin*, 23, 14-21.

Tulving, E. & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological review*, 80(5), 352-373.

VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.

Wang, Z. J. (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.

Wendel, J. (1997). Planning and second language narrative production. Unpublished doctoral dissertation. Temple University, Japan.

Wiener, M., Devoe, S., Rubinow, S., & Geller, J. (1972). Nonverbal behavior and nonverbal communication. *Psychological Review, 79,* 185-214.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14 (1), 85-106.

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

Appendices

Appendix 1 C-test

C-Test

This test will be scored anonymously. That is, you don't have to provide your name. But please honestly provide the results of the following exams if you have taken them. Thank you!

HKCEE English Exam (英國語文成績) _____

A-Level English Exam（英國語文成績） ____

Passage 1:

One cool autumn evening, Bob L., a young professional, returned home from a trip to the supermarket to find his computer gone. Gone! All so____ of cr____ thoughts ra____ through h_____ mind: H_____ it be_____ stolen? H_____ it be_____ kidnapped? H_____ searched h______ house f_____ a cl_____ until h_____ noticed a sm_____ piece o_______ printout pa______ stuck un______ a mag_______ on h______ refrigerator do______ . His heart sank as he read this simple message: CAN'T CONTINUE, FILE CLOZED, BYE!

Passage 2:

There is a third factor besides farming and herding in the spread of man-made deserts: deforestation. The progre_____ destruction o_____ the Th_____ World's st_____ of tr_____ is dama_____ not on_____ in d_____ regions: every_____ it occ_____ it c_____ accelerate t_____ decay o_____ the so_____ and red_____ its capa_____ to fe_____ people. It can reduce rainfall and lead to drought.

Passage 3:

There are certain things which no student can do without and others which may not be as necessary as you thought. It m_____ be wo_____ considering so_____ small hi_____. You m_____ find your_____ in ne_____ of elect_____ appliances su_____ as li_ ____ bulbs, adap_____ or pl_____ . These c_____ be obta_____ from ma_____ places. Gill i_____ a go_____ hardware sh_____ and try_____ to fi_____ it i_____ a chall_____ . It is hidden in a little alley leading off High Street called Wheatsheaf Yard.

Passage 4:

The private conscience of the leader – rather than his public responsibility – becomes the focal point of politics. Internal crit_____ – possession o_____, devotion t_____, and stan_____ up f_____ private prin_____ – become t_____ standards o_____ political judg_____ . Constituents disa_____ , and w_____ are le_____ with a poli_____ leader deter_____ policy o_____ the ba_____ of compa_____ with h_____ private princ_____ . From this perspective we can better understand why Goldwater voted against the Civil Rights Act of 1964.

Thank you very much. You think this test is (please tick)

Very easy____easy_____neutral______difficult________ very difficult______

Key:

Passage 1:

One cool autumn evening, Bob L., a young professional, returned home form a trip to the supermarket to find his computer gone. Gone! All sorts of crazy thoughts raced through his mind: Had it been stolen? Had it been kidnapped? He searched his house for a clue until he noticed a small piece of printout paper stuck under a magnet on his refrigerator door. His heart sank as he read this simple message: CAN'T CONTINUE, FILE CLOZED, BYE!

Passage 2:

There is a third factor besides farming and herding in the spread of man-made deserts: deforestation. The progressive destruction of the Third World's stock of trees is damaging not only in dry regions: everywhere it occurs it can accelerate the decay of the soil and reduce its capacity to feed people. It can reduce rainfall and lead to drought.

Passage 3:

There are certain things which no student can do without and others which may not be as necessary as you thought. It may be worth considering some small hints. You may find yourself in need of electrical appliances such as light bulbs, adaptors or plugs. These can be obtained from many places. Gill is a good hardware shop and trying to find it is a challenge. It is hidden in·a little alley leading off High Street called Wheatsheaf Yard.

Passage 4:

The private conscience of the leader – rather than his public responsibility – becomes the focal point of politics. Internal criteria – possession of, devotion to, and standing up for private principles – become the standards of political judgment. Constituents disappear, and we are left with a political leader determining policy on the basis of compatibility with his private principles. From this perspective we can better understand why Goldwater voted against the Civil Rights Act of 1964.

Source: Dornyer, Z and Katona, L (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, 9, 187-206.

Appendix 2 Survey of participant background

<table>
<tr><td colspan="6">Survey of Participant Background Information

All data will be used in this research only and will be discarded once the study finishes.</td></tr>
<tr><td rowspan="2">Chinese Name:</td><td rowspan="2">Gender:

M____ F___</td><td rowspan="2" colspan="2">Please tick as

appropriate:</td><td>Computer</td><td>Major___Minor___</td></tr>
<tr><td>Medicine</td><td>Major___Minor___</td></tr>
<tr><td colspan="2">Email:</td><td colspan="4">Phone:</td></tr>
<tr><td colspan="6">Your Overall result of : HKCEE_____ A-Level_____ IELTS_____

TOEFL______ Others (Please Specify)______

Your oral test result of: HKCEE______ A-Level______ IELTS______

TOEFL______ Others (Please Specify)_______</td></tr>
</table>

Appendix 3 Survey of Topic Familiarity

Survey of Topic Familiarity					
Chinese Name:	Major: Computer____ Medicine_______				
1: Totally ignorant (know nothing). 2: know a bit 3. A bit Familiar. 4. Quite familiar. 5:Very Familiar. **Please darken the circle under the appropriate number to indicate how familiar you are with each topic.**					
	1	**2**	**3**	**4**	**5**
Please describe in detail the general process of the infection of virus in a human body, and the possible consequences, the general procedure of dealing with a virus-infected person.	O	O	O	O	O
Please describe in detail the general process of the infection of virus in a computer, the possible consequences, and the general procedure of dealing with a virus-infected computer.	O	O	O	O	O

Appendix 4 A self-devised C-Test in pilot study 1

Passage 1:

London, is the largest city in Europe. It has held this title for over four hundred years and over 7,000,000 people call London, the capital of the United Kingdom, home. One in ten people, who live in the United Kingdom, live in London. 350,000 people travel each day into London to work. Due to London's location, it is a very dry place all year. Although it is dry, London is often very cloudy. On average, it rains mildly just about every other day.

(Source:http://www.teach-nology.com/worksheets/language_arts/reading_comp/elem/ver17/)

Passage 2:

One cool autumn evening, Bob L., a young professional, returned home form a trip to the supermarket to find his computer gone. Gone! All sorts of crazy thoughts raced through his mind: Had it been stolen? Had it been kidnapped? He searched his house for a clue until he noticed a small piece of printout paper stuck under a magnet on his refrigerator door. His heart sank as he read this simple message: CAN'T CONTINUE, FILE CLOZED, BYE!

Borrowed from: Dornyer and Katona (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, 9, 187-206.

Appendix 5 Consent form for participants

Second Language Speaking Task Performance Research

Consent Form

Thank you very much for participating in the research conducted by Mr. Gavin Bei Xiaoyue, a Ph.D. student at the English Department at CUHK, to investigate second language speaking task performance.

As a research participant, you are entitled to the following definite rights:

1. Your answers to any questions will not have any influence on your class performance or your standing at the university.
2. You may withdraw from the study at any time and refuse to answer any particular question.
3. You may contact the researcher at any time to give your reactions and comments about the study.
4. All data collected, including your personal information and task performance, will be kept strictly confidential and will be available only to the researcher and his supervisor.
5. Data collected from you may be made part of the final research thesis, but please be reassured that *under no circumstances will your real name be included in the thesis*.
6. Your personal information will be destroyed once the research is finished.
7. If requested, the research will inform you the results when the study is completed.
6. Upon the completion of your participation, your will be given $ 50 honorarium.

I have carefully read the above and ☐ agree / ☐ disagree to participate in the study.

Full name:____________________

Signature:____________________

Date:________________________

Appendix 6 Eight sample coded files from the eight cells in the experiment

Note: please refer to section 3.7 'Data processing and coding scheme' for a detailed description of the coding scheme for the following samples.

Sample 1 (Medicine major, unplanned, unfamiliar)

*CYF: in fact I don't really know about it because I very weak in computer .

%mor: prep|in n|fact pro|I v|don't re#n|ally v|know prep|about pro|it conj:subor|because pro|I adv:int|very adj|weak prep|in n|computer .

%snd: <01.37.78><01.43.28>

%CYF: er (0.56) in fact I don't really know about it errfr ::: because I very weak in computer . err_m_s :;:a |

*CYF: I think it's some virus may affect the programs in the computers

%mor: pro|I v|think pro|it's qn|some n|virus aux|may v|affect det|the n|program-PL prep|in det|the n|computer-PL

%snd: <01.43.53><01.51.28>

%CYF: er I think errfr ::; {it's} # (0.40) er some virus may affect the programs er in the computers . err_m_m ::: |

*CYF: and make them can't function well so that when we are using these programs it doesn't work .

%mor: conj:coo|and v|make pro|them v|can't n|function adv|well conj:subor|so rel|that conj:subor|when pro|we aux|be&PRES part|use-PROG det|these v|program-3S pro|it v|doesn't n|work .

%snd: <01.51.70><02.00.39>

%CYF: and {make them ca@} * er make them can't function well err_m_s ::: so that er when we are using these programs errfr :;:b (0.45) er it doesn't work . errfr :;:a |

*CYF: and carry some virus and may affect, is it, the quality of the file .

%mor: conj:coo|and v|carry qn|some n|virus conj:coo|and aux|may v|affect aux|be&3S pro|it det|the n|quality prep|of det|the n|file .

%snd: <02.00.57><02.12.03>

%CYF: {And} * and (0.70) carry (0.48) er some (0.44) virus err_m_m ::: and may affect (0.41) er {the quality} * is it , the quality {of} * (0.82) of the file . errfr ::: |

*CYF: I don't really know it, but .

%mor: pro|I v|don't re#n|ally v|know pro|it conj:coo|but .

%snd: <02.12.34><02.13.48>

%CYF: I don't really know it , {but} # . err_a_d ::: |

*CYF: and then some virus may make the computer bad .

%mor: conj:coo|and adv:tem|then qn|some n|virus aux|may v|make det|the n|computer adj|bad .

%snd: <02.13.82><02.28.22>

%CYF: {and} * (0.54) and then some virus {may (0.47) er cause * (0.72) er er may cause * cause our computer} ~ (1.03) may make er (0.54) the computer (0.77) bad . err_a_l ::: |

*CYF: I mean we can't open the computer .

%mor: pro|I v|mean pro|we v|can't v|open det|the n|computer .

%snd: <02.28.92><02.30.94>

%CYF: I mean errfr ::: we can't open the computer . err_s_l :;:a |

*CYF: we can't turn on the computer .

%mor: pron|we v|can't v|turn prep|on det|the n|computer .

%snd: <02.31.05><02.33.26>

%CYF: we can't turn on the computer , and_pfp . errfr ::: |

*CYF: all programs become a mess .

%mor: qn|all n|program-PL v|become det|a n|mess .

%snd: <02.33.77><02.35.38>

%CYF: all programs become a mess . errfr ::: |

*CYF: and then at the same time we may forward the virus to other people through our online system, for example, MSN .

%mor: conj:coo|and adv:tem|then prep|at det|the adj|same n|time pro|we aux|may v|forward det|the n|virus prep|to qn|other v|people prep|through pro:poss:det|our adj|online n|system prep|for n|example prop:n|MSN .

%snd: <02.35.82><02.43.69>

%CYF: and then at the same time we may forward the virus to other people through our online system , for example , MSN . errfr ::: |

*CYF: the virus will makes our online programs, for example, MSN to automatically to send this virus out to other people and at the same time affect the other people's computer .

%mor: det|the n|virus aux|will v|make-3S pro:poss:det|our adj|online n|program-PL prep|for n|example pro:n|MSN prep|to adv:adj|automatic-LY inf|to v|send det|this n|virus n|out prep|to qn|other n|person&PL conj:coo|and prep|at det|the adj|same n|time v|affect det|the qn|other n|people's n|computer .

%snd: <02.44.29><02.58.93>

%CYF: er er the virus will (0.47) makes our (0.44) online programs , for example , MSN err_m_m ::: to automatically to send this virus out to other people err_m_s :;;a and at the same time affect {the} * the other people's (0.42) computer .

err_m_m err_m_r ::: |

*CYF: and some virus may be used to steal some personal information from our computers, for example, our login ID or some passwords, or some online banking information, etc .

%mor: conj:coo|and qn|some n|virus aux|may v|be adv|used inf|to v|steal qn|some adj|personal n|information prep|from pro:poss:det|our n|computer-PL prep|for n|example pro:poss:det|our n|login n|ID conj:coo|or qn|some n|passwords conj:coo|or qn|some adj|online part|bank-PROG n|information adv|etc .

%snd: <02.59.22><03.19.25>

%CYF: and some virus may be used err_m_m ::: (0.76) to er steal some personal information from our computers , for example , our login ID or (0.91) some passwords , or some online banking er information , etc . errfr :;;a |

*CYF: I think that the virus is very dangerous in fact because nowadays you know everyone has computers .

%mor: pro|I v|think pro:dem|that det|the n|virus v|be&3S adv:int|very adj|dangerous prep|in n|fact conj:subor|because adv:tem|nowadays pro|you v|know pro:indef|everyone v|have&3S n|computer-PL .

%snd: <03.19.96><03.31.65>

%CYF: {and that} # (0.46) I think errfr ::; that the virus is very dangerous in fact errfr ::: because nowadays er (0.50) er you_know_pfp {everyone has} * er , everyone has computers . err_m_m :;:a |

*CYF: they may put their personal informations or personal things in the computer

%mor: pro|they aux|may v|put aux|may prep|of pro:poss:det|their adj|personal n|information-PL conj:coo|or adj|personal n|thing-PL prep|in det|the n|computer

%snd: <03.31.69><03.38.51>

%CYF: they may put {may} # (0.40) of their personal informations or personal things in the computer . err_m_s ::: |

*CYF: and then at the same time they will use the computers to do many things .

%mor: conj:coo|and adv:tem|then prep|at det|the adj|same n|time pro|they aux|will v|use det|the n|computer-PL inf|to v|do qn|many n|thing-PL .

%snd: <03.38.83><03.43.30>

%CYF: and then at the same time they will use the computers to do many things . errfr ::: |

*CYF: if the computers is infected by the virus, then they may not have normal life .

%mor: conj:subor|if det|the n|computer-PL aux|be&3S part|infect-PERF prep|by det|the n|virus adv:tem|then pro|they aux|may neg|not v|have adv|normal n|life .

%snd: <03.43.40><03.54.55>

%CYF: If the computers (0.51) is infected by the virus err_m_m err_m_s :;:b then {they may not} * (0.48) they may not er (0.82) have normal life (0.70) er yeah er . err_m_s ::: |

*CYF: and it is not safe .

%mor: conj:coo|and pro|it v|be&3S neg|not adj|safe .

%snd: <03.55.62><04.00.60>

%CYF: {And (0.94) they're} ~ and (0.78) it is not safe . errfr ::: |

*CYF: and now I am going to talk about the procedure to deal with the virus infected computers .

%mor: conj:coo|and adv|now pro|I aux|be&1S part|go-PROG inf|to v|talk prep|about det|the n|procedure prep|to n|deal prep|with det|the n|virus v|infect-PAST n|computer-PL .

%snd: <04.02.55><04.12.32>

%CYF: and (1.42) er now XX I am going to talk about the procedure errfr ::: to deal with the virus infected (1.03) computers . errfr :;;a |

*CYF: I think there is some programs online that we can download .

%mor: pro|I v|think adv:loc|there v|be&3S qn|some n|program-PL adj|online pro:dem|that pro|we aux|can v|download .

%snd: <04.12.35><04.15.52>

%CYF: I think errfr :::; there is some programs online err_m_s ::: that we can download . errfr :;:a |

*CYF: Norton , I don't know , this Norton may be one of er these programs .

%mor: n:prop|Norton pro|I v|don't v|know pro:dem|this n:prop|Norton aux|may v|be pro:indef|one prep|of fil|er det|these n|program-PL .

%snd: <04.15.96><04.21.19>

%CYF: Norton # I don't know err_a_d ::: this Norton may be one of (0.57) er these programs . errfr ::: |

*CYF: and we can download them and then we run that program and then it will help us to clear the virus automatically .

%mor: conj:coo|and pro|we aux|can v|download pro|them conj:coo|and adv:tem|then pro|we v|run det|that n|program conj:coo|and adv:tem|then pro|it aux|will v|help pro|us inf|to v|clear det|the n|virus adv:adj|automatic-LY .

%snd: <04.21.24><04.28.02>

%CYF: and we can download them errfr ::: and then we run that program err_m_r ::: and then (0.48) it will help us errfr ::: to clear the virus automatically . errfr :;;a |

*CYF: if that program doesn't work we may ask some professions .

%mor: conj:subor|if det|that n|program ?|doesn't n|work pro|we aux|may v|ask qn|some n|profession-PL .

%snd: <04.28.39><04.32.76>

%CYF: er if {that} * (0.41) er that program doesn't work errfr :;:b we may ask some professions . err_s_l ::: |

*CYF: we may employ them and ask them to help us to settle that problem .

%mor: pro|we aux|may v|employ pro|them conj:coo|and v|ask pro|them inf|to v|help pro|us inf|to v|settle det|that n|problem .

%snd: <04.33.22><04.39.21>

%CYF: yeah_pfp we may employ them errfr ::: and ask them errfr ::: to help us errfr :;;a to settle that problem . errfr :;;a |

*CYF: I think that's all .

%mor: pro|I v|think pro|that's qn|all .

%snd: <04.40.34><04.41.03>

%CYF: I think errfr ::; that's all . errfr ::: |

@End

Sample 2 (Medicine major, unplanned, familiar)

*KMY: we cannot classify it as living or non-living because you can say it is living when it's attacked to body cell .

%mor: pro|we aux|can~neg|not v|classify pro|it prep|as part|live-PROG conj:coo|or adj|non-living conj:subor|because pro|you aux|can v|say pro|it aux|be&3S part|live-PROG adv:wh|when pro|it's v|attack-PAST prep|to n|body n|cell .

%snd: <01.13.24><01.23.73>

%KMY: er {virus is actually} # er we cannot classify it as living or non-living errfr ::: because {it is} # you can say errfr ::: it is living errfr :;:a when it's attacked to body cell . err_m_m err_s_p :;:a |

*KMY: actually how how it's work .

%mor: adv|actually adv:wh|how adv:wh|how pro|it's n|work .

%snd: <01.24.18><01.27.41>

%KMY: actually_pfp {how it's work} * {because} # {how} * how it's work . err_m_s ::: |

*KMY: virus got a fragile mechanism .

%mor: n|virus part|get&PERF det|a adj|fragile n|mechanism .

%snd: <01.27.73><01.30.25>

%KMY: er virus got a fragile mechanism . err_m_s ::: |

*KMY: it can infect the body, as for example, human body cell .

%mor: pro|it aux|can v|infect det|the n|body conj:subor|as prep|for n|example adj|human n|body n|cell .

%snd: <01.30.70><01.34.75>

%KMY: it can er infect the body as for example , human body cell . err_m_m ::: |

*KMY: it can infect it and just er inject their own DNA or circulate DNA into our cell .

%mor: pro|it aux|can v|infect pro|it conj:coo|and adv:int|just fil|er v|inject pro:poss:det|their adj|own n:prop|DNA conj:coo|or v|circulate n:prop|DNA prep|into pro:poss:det|our n|cell .

%snd: <01.35.16><01.36.01>

%KMY: it can infect it err_m_r ::: and just er inject {the} ~ their own DNA errfr ::: {or circulate} * (0.40) er or circulate DNA into our cell . err_m_s err_m_m ::: |

*KMY: and this circulate DNA can insert into our own DNA genome so that our DNA may be controlled by the virus DNA .

%mor: conj:coo|and det|this v|circulate ?|DNA aux|can v|insert prep|into pro:poss:det|our adj|own prop:n|DNA n|genome conj:subor|so pro:dem|that pro:poss:det|our prop:n|DNA aux|may v|be part|control-PERF prep|by det|the n|virus prop:n|DNA .

%snd: <01.45.58><01.55.05>

%KMY: and this circulate DNA can insert into our own DNA genome err_m_s ::: (0.40) so that our DNA {will} rpl (0.42) may be controlled by the virus DNA .

errfr :;:a |

*KMY: and the immune system will be control and destroy . .

%mor: conj:coo|and det|the adj|immune n|system aux|will v|be n|control conj:coo|and v|destroy . .

%snd: <01.55.29><02.01.44>

%KMY: and {the} * the er immune system will be (0.44) er control and destroy . err_m_s ::: |

*KMY: and how destroy is that the virus DNA will control our DNA and to express some gene and express some protein that will destroy our own cells .

%mor: conj:coo|and adv:wh|how v|destroy v|be&3S pro:dem|that det|the n|virus prop:n|DNA v|will v|control pro:poss:det|our prop:n|DNA conj:coo|and prep|to adj|express qn|some n|gene conj:coo|and adj|express qn|some n|protein pro:dem|that aux|will v|destroy pro:poss:det|our adj|own n|cell-PL .

%snd: <02.01.83><02.17.48>

%KMY: and how destroy is err_s_s ::: that {they} rpl the {virus} * virus DNA will er er control our DNA errfr :;:a and to express some gene err_m_m :;:a and express some protein errfr :;:a that will destroy our own cells . errfr :;:a |

*KMY: and this can lead to the breakage or the die of the cell .

%mor: conj:coo|and pro:dem|this aux|can v|lead prep|to det|the n|breakage conj:coo|or det|the n|die prep|of det|the n|cell .

%snd: <02.17.62><02.22.37>

%KMY: and this can lead to {the leak} ~ the breakage or the die of the cell . err_m_l ::: |

*KMY: and the consequence is that many cells could die .

%mor: conj:coo|and det|the n|consequence v|be&3S adv:int|that qn|many n|cell-PL aux|could v|die .

%snd: <02.23.60><02.27.04>

%KMY: and the consequence is that many cells could die . errfr ::: |

*KMY: and some cells maybe not die, but they just control by the virus .

%mor: conj:coo|and qn|some n|cell-PL adv|maybe neg|not v|die conj:coo|but pro|they adv:int|just v|control prep|by det|the n|virus .

%snd: <02.27.33><02.32.00>

%KMY: and some cells er maybe not die , err_m_s ::: but they just control by the virus . err_m_s ::: |

*KMY: so they will work for the virus and attack other part of the body .

%mor: co|so pro|they aux|will v|work prep|for det|the n|virus conj:coo|and v|attack qn|other n|part prep|of det|the n|body .

%snd: <02.32.29><02.37.50>

%KMY: {show they} * so they will work for the virus errfr ::: and attack other part of the body . err_m_m ::: |

*KMY: and attack other part of the immune system .

%mor: conj:coo|and v|attack qn|other n|part prep|of det|the adj|immune n|system .

%snd: <02.38.11><02.40.06>

%KMY: and attack other part of the immune system . err_m_m ::: |

*KMY: and the very important example is that the HIV .

%mor: conj:coo|and det|the adv:int|very adj|important n|example v|be&3S pro:dem|that det|the ?|HIV .

%snd: <02.40.49><02.47.05>

%KMY:and the {very er im@} ~ very important er (0.46) example is that the HIV . err_m_s ::: |

*KMY: it works like this and it attack T cells of the bodies .

%mor: n:prop|It v|work-3S prep|like det|this conj:coo|and pro|it n|attack n:prop|T n|cell-PL prep|of det|the n|body-PL .

%snd: <02.47.60><02.50.81>

%KMY: it works like this errfr ::: and it attack T cells of the bodies . err_m_s ::: |

*KMY: the T cell is one of the immune system cell in our body .

%mor: det|the n:prop|T n|cell v|be&3S det:num|one prep|of det|the adj|immune n|system n|cell adv:loc|in pro:poss:det|our n|body .

%snd: <02.51.24><02.54.38>

%KMY: the T cell is one of the immune system cell in our body . err_m_m ::: |

*KMY: virus can can group into many type .

%mor: n|virus n|can aux|can v|group prep|into qn|many n|type .

%snd: <02.54.92><03.02.89>

%KMY: er actually_pfp {virus er des@} ~ virus {can} * er can group into many type . err_m_m err_m_s ::: |

*KMY: some may not be killed .

%mor: qn|some aux|may neg|not v|be part|kill-PERF .

%snd: <03.03.09><03.06.18>

%KMY: {some maybe very} ~ some may not be killed .errfr ::: |

*KMY: some can be killed .

%mor: qn|some aux|can v|be v|killed .

%snd: <03.06.73><03.09.51>

%KMY: er some {may be} ~ can be killed er . errfr ::: |

*KMY: for example some influence can be cured because we have anti-biotics or something like that .

%mor: n:prop|For n|example qn|some n|influence aux|can v|be part|cure-PERF conj:subor|because pro|we v|have n|anti-biotics conj:coo|or pro:indef|something prep|like pro:dem|that .

%snd: <03.10.10><03.20.12>

%KMY: for example er {some er influence} * (0.41) some influence can be cured err_s_l ::: because (0.40) er (0.50) er we have anti-biotics or something like that . errfr :;:a |

*KMY: but some do not have the vaccine and cannot cure you .

%mor: conj:coo|but qn|some aux|do neg|not v|have det|the n|vaccine conj:coo|and aux|can~neg|not v|cure pro|you .

%snd: <03.20.49><03.23.90>

%KMY: but some do not have the vaccine err_m_r ::: and cannot cure you . err_m_r ::: |

*KMY: and nowadays the many cocktails treatment for the some disease that cause by virus infection er such as HIVs .

%mor: conj:coo|and adv:tem|nowadays det|the qn|many n|cocktail-PL n|treatment prep|for det|the qn|some dis#v|ease pro:dem|that n|cause prep|by n|virus n|infection fil|er qn|such prep|as prop:n|HIVs .

%snd: <03.24.55><03.38.68>

%KMY: er (1.37) and nowadays the many cocktails er treatment for the some disease err_m_m ::: that {cause by virus} ~ and cause by {virus in@} * virus infection er such as HIVs . err_m_m err_m_s :;:a |

*KMY: it is still in the testing process but it may work in the future .

%mor: pro|it v|be&3S adj|still adv:loc|in det|the part|test-PROG n|process conj:coo|but pro|it aux|may v|work adv:loc|in det|the adj|future .

%snd: <03.39.18><03.43.84>

%KMY: er it is still in the testing process errfr ::: but it may work in the future . errfr ::: |

*KMY: yeah, that's all .

%mor: adv|yeah pro|that's qn|all .

%snd: <03.44.35><03.45.00>

%KMY: Yeah, that's all . errfr ::: |

@End

Sample 3 (Medicine major, planned, unfamiliar)

*MHY: we will get sick because of virus , but computer can also get sicks because of the IT virus .

%mor: pro|we aux|will v|get adj|sick conj:subor|because prep|of n|virus conj:coo|but n|computer n|can adv|also v|get n|sick-PL conj:subor|because prep|of det|the n:prop|IT n|virus .

%snd: <10.02.26><10.09.70>

%MHY: we will get sick because of virus , err_m_m ::: but er computer can also get sicks because of the {virus} ~ IT virus . err_m_m ::: |

*MHY: some of the examples like Tro# Horse, which is hidden in some of the files in the Internet .

%mor: qn|some prep|of det|the n|example-PL co|like Tro#n:prop n:prop|Horse rel|which aux|be&3S part|hide&PERF prep|in qn|some prep|of det|the n|file-PL prep|in det|the n:prop|Internet .

%snd: <10.10.18><10.19.00>

%MHY: er some of the (0.41) er examples like Tro@ Horse err_m_l err_m_s ::: (0.45) er which is hidden in {some of} * er some of the files in the Internet . err_m_l :;:a |

*MHY: when you download the files, the Trojan Horse virus will be downloaded together .

%mor: conj:subor|when pro|you v|download det|the n|file-PL det|the n:prop|Trojan n:prop|Horse n|virus aux|will v|be part|download-PERF adv|together .

%snd: <10.19.01><10.33.16>

%MHY: when you download er the files errfr :;:b {you} * {you} # er {the} * {the Horse} ~ the Trojan Horse (1.00) virus {can} rpl er (0.64) mmm (1.10) er will be downloaded er together . errfr ::: |

*MHY: and this file will be hidden in some of the hidden place in the computer .

%mor: conj:coo|and det|this n|file aux|will v|be part|hide&PERF prep|in qn|some prep|of det|the part|hide&PERF n|place prep|in det|the n|computer .

%snd: <10.33.49><10.44.65>

%MHY: {so} # er (1.06) {and} * and this er (0.61) file will be hidden {in * er in some of the place} ~ (0.61) er {in} * in some of the hidden place in the computer . err_m_m ::: |

*MHY: or may even affect the function of the computer .

%mor: conj:coo|or aux|may v|even v|affect det|the n|function prep|of det|the n|computer .

%snd: <10.45.17><10.48.65>

%MHY: or may even er affect the function of the computer . err_m_m ::: |

*MHY: apart from the Trojan Horse file, there are a lot of types of the virus in the Internet , which can attack our computer .

%mor: adv|apart prep|from det|the n:prop|Trojan n:prop|Horse v|file adv:loc|there v|be&PRES det|a n|lot prep|of n|type-PL prep|of det|the n|virus prep|in det|the n:prop|Internet rel|which aux|can v|attack pro:poss:det|our n|computer .

%snd: <10.49.07><11.12.06>

%MHY: er {other virus are} # (0.88) er apart from the Trojan Horse file er {there is some} ~ (0.67) there are {some of the (0.60) virus} rpl er (0.83) er {a lot} * (1.37) a lot of er types of the virus in the Internet , err_m_m err_m_l ::: {which} * er (1.49) which can attack our computer . err_m_m :;:a |

*MHY: some of the virus can also be spread from other infected person .

%mor: qn|some prep|of det|the n|virus n|can adv|also v|be n|spread prep|from qn|other v|infect-PAST n|person .

%snd: <11.12.98><11.22.25>

%MHY: mmm (0.83) some of the virus {can also be er infected} rpl er (0.43) {can al@} * can also be spread from other infected person . err_m_m ::: |

*MHY: once you are infected, that means your computer is infected, they will stay in some of the important files which can not be able to restore or deleted because if you delete the file or restore it, it may affect the normal function of your computer, just like er system32 EXE .

%mor: adv|once pro|you aux|be&PRES part|infect-PERF pro:dem|that v|mean-3S pro:poss:det|your n|computer aux|be&3S part|infect-PERF pro|they aux|will v|stay prep|in qn|some prep|of det|the adj|important n|file-PL rel|which n|can neg|not v|be adj|able inf|to re#n|store conj:coo|or part|delete-PERF conj:subor|because conj:subor|if pro|you v|delete det|the n|file conj:coo|or v|restore pro|it pro|it aux|may v|affect det|the adj|normal n|function prep|of pro:poss:det|your n|computer adv:int|just prep|like fil|er ?|system32 ?|EXE .

%snd: <11.22.56><11.55.34>

%MHY: mmm once you are infected errfr :;:b (0.47) er that means errfr :::; your computer {is er in@} * is infected errfr :;:a er they will stay in some of the important files err_m_r ::: which can not be able to restore or deleted err_m_l :;:a because er if you delete the file errfr :;:b or restore it errfr :;:b er it may er affect er {the function of} ~ the normal function {of your} * (1.22) of your computer , {just like the er system32} * (1.56) just like er system32 er EXE mmm . err_m_l err_m_m ::: |

*MHY: in fact, the virus just like influenza because once you are infected, you may spread to the others which is not deliberate .

%mor: prep|in n|fact det|the n|virus adv:int|just prep|like n|influenza conj:subor|because adv|once pro|you aux|be&PRES v|infect-PAST pro|you aux|may n|spread prep|to det|the pro:indef|other-PL rel|which v|be&3S neg|not v|liberate .

%snd: <11.56.80><12.14.54>

%MHY: in fact {infected} # (1.16) er the virus just like influenza because er once you are infected errfr :;:b {you may} * er (1.24) you may er spread to the others er {without} # er err_a_l ::: (2.46) which is not er deliberate . errfr :;:a |

*MHY: and you may send infected files through the MSN automatically .

%mor: conj:coo|and pro|you aux|may v|send part|infect-PERF n|file-PL prep|through det|the n:prop|MSN adv:adj|automatic-LY .

%snd: <12.15.71><12.21.66>

%MHY: and$ you may send {infect} ~ infected files through the MSN automatically . errfr ::: |

*MHY: and if your computer is infected, your computer may not have a normal functioning or may even restart .

%mor: conj:coo|and conj:subor|if pro:poss:det|your n|computer aux|be&3S part|infect-PERF pro:poss:det|your n|computer aux|may neg|not v|have det|a adj|normal part|function-PROG conj:coo|or aux|may adj|even re#n|start .

%snd: <12.21.76><12.33.89>

%MHY: and er (0.48) if your er computer is infected errfr :;:b your computer may$ not have a normal functioning err_m_l ::: or may even restart er . errfr ::: |

*MHY: and .

%mor: conj:coo|and .

%snd: <12.35.07><12.35.43>

%MHY: and_pfp . |

*MHY: and you may also not be able to open some of the files just like the Internet .

%mor: conj:coo|and pro|you aux|may adv|also neg|not v|be adj|able inf|to v|open qn|some prep|of det|the n|file-PL adv:int|just prep|like det|the n:prop|Internet .

%snd: <12.37.28><12.45.71>

%MHY: and {you may also$ not} * (0.58) you may also not be able to open some of the files er just like er the Internet . err_m_p err_m_r ::: |

*MHY: in order to fix the problem you we can found the IT support from the website just like PCCiline website or from your university or from some of the IT friends .

%mor: prep|in n|order inf|to v|fix det|the n|problem pro|you pro|we aux|can v|find&PAST det|the ?|IT n|support prep|from det|the n|website adv:int|just prep|like n:prop|PCCiline n|website conj:coo|or prep|from pro:poss:det|your n|university conj:coo|or prep|from qn|some prep|of det|the n|IT n|friend-PL .

%snd: <12.46.96><13.03.50>

%MHY: er (1.19) in order to fix the problem errfr :;;b er {you} rpl (0.80) we can found the IT support from the website er (0.57) just like PCCiline website or from your university or from some of the er IT friends . err_m_s ::: |

*MHY: and other methods to fix the problem can be done by deleting and restoring the files .

%mor: conj:coo|and qn|other n|method-PL inf|to v|fix det|the n|problem aux|can v|be part|do&PERF prep|by part|delete-PROG conj:coo|and part|restore PROG det|the n|file-PL .

%snd: <13.04.02><13.14.95>

%MHY: and (1.21) to fix the problem errfr :;;b other methods {can} * (0.58) can be done err_m_s ::: by {deleted} ~ (0.54) er deleting (0.43) and restoring the files . errfr :;;a |

*MHY: and after deleting and restoring the files, it is necessary to quarantine the virus to prevent from the same infection again .

%mor: conj:coo|and prep|after part|delete-PROG conj:coo|and part|restore-PROG det|the n|file-PL pro|it v|be&3S adj|necessary inf|to v|quarantine det|the n|virus prep|to n|vent prep|from det|the adj|same n|infection adv|again .

%snd: <13.15.61><13.29.92>

%MHY: and after deleting and restoring the files errfr :;;b er it is necessary errfr ::: to quarantine the virus errfr :;;a (0.45) er (0.55) {so that} # to prevent from er (1.11) {another infect@} ~ the same infection again . err_m_s :;;a |

*MHY: the last thing we have to do is to update the anti-virus software in order to get rid of some of the updated new virus .

%mor: det|the adj|last n|thing pro|we v|have inf|to v|do v|be&3S prep|to adj|date det|the adj|anti-virus n|software prep|in n|order inf|to v|get ?|rid prep|of qn|some prep|of det|the part|date-PERF adj|new n|virus .

%snd: <13.30.57><13.42.77>

%MHY: er we have to do errfr :;:m the last thing is to update the anti-virus software errfr ::: in order {to} * er (1.10) to get rid of some of the (0.47) er updated {new part} rpl new virus . err_m_m :;;a |

@End

Sample 4 (Medicine major, planned, Familiar)

JCC: upon contact with the irus, the virus will be internalized into the human body .

%mor: prep|upon n|contact prep|with det|the n|virus det|the n|virus aux|will v|be part|internalize-PERF prep|into det|the n|human n|body .

%snd: <01.11.63><01.23.21>

%JCC: well_pfp (0.55) mmm (1.10) {upo@} * mmm upon contact with the virus , (0.50) er the virus will be internalized into the human body . errfr ::: |

*JCC: and the virus will attach to different types of human cell depending on its specificity .

%mor: conj:coo|and det|the n|virus aux|will v|attach prep|to adj|different n|type-PL prep|of adj|human n|cell part|depend-PROG prep|on pro:poss:det|its n|specificity.

%snd: <01.23.46><01.30.55>

%JCC: and the virus will attach to different types of human cell err_m_m ::: depending on its specificity . errfr :;;a |

*JCC: and then it will be internalized into a human cell when it replicates, using the protein in the human cell .

%mor: conj:coo|and adv:tem|then pro|it aux|will v|be part|internalize-PERF prep|into det|a n|human n|cell conj:subor|when pro|it ?|replicates part|use-PROG det|the n|protein prep|in det|the n|human n|cell .

%snd: <01.31.06><01.41.32>

%JCC: and then it will be internalized into a human cell errfr ::: when it replicates errfr :;:a (1.41) using the er protein in the human cell . errfr :;;a |

*JCC: after replication, it will cause lyses of the cell where the cell will burst and the viruses will be released into the blood stream .

%mor: prep|after n|replication pro|it aux|will v|cause ?|lyses prep|of det|the n|cell adv:wh|where det|the n|cell v|will n|burst conj:coo|and det|the n|virus-PL v|will v|be part|release-PERF prep|into det|the n|blood n|stream.

%snd: <01.41.81><01.53.01>

%JCC: after replication (0.50) {it} * er it will cause lyses of the cell errfr ::: where the cell will burst errfr :;:a and the viruses will be released into (0.48) the blood stream . errfr :;:a |

*JCC: through the blood stream, it will then infect other human cells .

%mor: prep|through det|the n|blood v|stream pro|it v|will adv:tem|then v|infect qn|other adj|human n|cell-PL .

%snd: <01.54.12><02.01.04>

%JCC: through the blood stream {it will then} * (1.03) er it will then infect other human cells . errfr ::: |

*JCC: .

%mor: fil|mmm .

%snd: <02.03.11><02.03.50>

%JCC: mmm . |

*JCC: XX infection .

%mor: n|XX n|infection .

%snd: <02.05.34><02.06.53>

%JCC: XX infection . |

*JCC: and after the human body is infected with the virus, firstly inflammation will occur where the eosinophils and other types of white blood cell in the human body will cause the inflammation stage .

%mor: conj:coo|and prep|after det|the n|human n|body aux|be&3S part|infect-PERF prep|with det|the n|virus adv:adj|first-LY n|inflammation aux|will v|occur adv:wh|where det|the n|eosinophils conj:coo|and qn|other n|type-PL prep|of adj|white n|blood n|cell prep|in det|the n|human n|body aux|will v|cause det|the n|inflammation n|stage .

%snd: <02.06.79><02.41.14>

%JCC: and er (0.54) {con@} # well_pfp after the human body is infected with the virus errfr :;:b (0.59) er (0.91) firstly inflammation will occur errfr ::: (0.44) where the er eosinophils and other types of white blood cell in the human body (0.53) er {will cause (0.44) the inflama@} ~ will cause the inflammation stage . err_m_m :;:a |

*JCC: after that, interferon will be released, which will kill the infected cell and prevent further replication of the virus .

%mor: prep|after pro:dem|that ?|interferon aux|will v|be part|release-PERF rel|which aux|will v|kill det|the part|infect-PERF n|cell conj:coo|and adj|vent v|further n|replication prep|of det|the n|virus .

%snd: <02.31.82><02.42.53>

%JCC: after that , er interferon will be released errfr ::: (0.53) which will kill the infected cell errfr :;:a and prevent further (0.46) replication of the virus . errfr :;:a |

*JCC: the complement system will also be activated .

%mor: det|the n|complement n|system v|will adv|also v|be part|activate-PERF .

%snd: <02.43.19><02.46.47>

%JCC: the complement system will also be activated . errfr ::: |

*JCC: after which natural killer cells will be activated and they will kill all the rest of the virus .

%mor: prep|after rel|which adj|natural n:v|kill-AGT n|cell-PL v|will v|be v|activate-PAST conj:coo|and pro|they aux|will v|kill qn|all det|the n|rest prep|of det|the n|virus .

%snd: <02.47.46><02.55.23>

%JCC: after which natural killer cells will be activated errfr ::: and they will kill (0.57) all the rest of the virus . err_m_m ::: |

*JCC: when that fails, the cell mediated immunity of the human body will be activated .

%mor: adv:wh|when pro:dem|that v|fail-3S det|the n|cell v|mediate-PAST n|immunity prep|of det|the n|human n|body aux|will v|be part|activate-PERF .

%snd: <02.56.70><03.05.28>

%JCC: mmm when that fails errfr :;:b mmm the cell mediated er immunity of the human body will be activated . errfr ::: |

*JCC: and the immuno-globin will be released to kill the virus .

%mor: conj:coo|and det|the ?|immuno-globin aux|will v|be v|release-PAST inf|to v|kill det|the n|virus .

%snd: <03.05.91><03.11.55>

%JCC: and the immuno-globin will be released errfr ::: to kill the er virus . errfr :;;a |

*JCC: the consequence of which is tissue destruction .

%mor: det|the n|consequence prep|of rel|which v|be&3S n|tissue n|destruction.

%snd: <03.13.46><03.17.16>

%JCC: the consequence of which is tissue destruction . errfr :;:a |

*JCC: and as a result the human body will develop a fever .

%mor: conj:coo|and prep|as det|a n|result det|the n|human n|body aux|will v|develop det|a n|fever .

%snd: <03.17.81><03.22.24>

%JCC: and as a result er the human body will develop a fever . errfr ::: |

*JCC: for example when a human body is infected with a rhino virus, which the human will develop symptoms like a common flu .

%mor: prep|for n|example adv:wh|when det|a n|human n|body aux|be&3S part|infect-PERF prep|with det|a n|rhino n|virus rel|which det|the n|human aux|will v|develop n|symptom-PL v|like det|a adj|common n|flu .

%snd: <03.23.51><03.45.60>

%JCC: mmm (0.95) {other conce@} # er {for exam@} * well_pfp for example er (0.76) when a human body is infected {with a rhino} * (0.51) with a rhino virus errfr :;:b (0.59) mmm {which} # (0.56) er the human will develop symptoms er like a common flu . errfr ::: |

*JCC: and that includes fever, soar throat, fatigue, dehydration which is a result of the fever .

%mor: conj:coo|and pro:dem|that v|include-3S n|fever v|soar n|throat n|fatigue n|dehydration rel|which v|be&3S det|a n|result prep|of det|the n|fever.

%snd: <03.46.52><03.57.61>

%JCC: and that includes fever , soar throat , fatigue , (0.47) er dehydration errfr ::: (0.62) which is er a result of the fever . errfr :;:a |

*JCC: and more serious consequences are cancer .

%mor: conj:coo|and qn|more adj|serious n|consequence-PL v|be&PRES n|cancer .

%snd: <03.58.41><04.05.41>

%JCC: and more serious consequences mmm (1.03) mmm are cancer . err_m_s ::: |

*JCC: in hepatitis B and C, after the liver cells have been infected with virus, the virus will in its replication process alter the DNA of the human liver cells .

%mor: prep|in n|hepatitis n:prop|B conj:coo|and n:prop|C prep|after det|the n|liver n|cell-PL aux|have part|be&PERF v|infect-PAST prep|with n|virus det|the n|virus aux|will prep|in pro:poss:det|its n|replication n|process v|alter det|the ?|DNA prep|of det|the n|human n|liver n|cell-PL .

%snd: <04.06.12><04.23.57>

%JCC: mmm in hepatitis B and C (0.53) mmm after the liver cells {hav@} *

have been er infected with virus errfr :;:b (0.54) the virus will in its replication process (0.61) alter the DNA {of the liver} rpl of the human liver cells . errfr ::: |

*JCC: and as a result, a mutation of the liver cell will occur and cancer will be eventually be developed .

%mor: conj:coo|and prep|as det|a n|result det|a ?|mutation prep|of det|the n|liver n|cell aux|will v|occur conj:coo|and n|cancer aux|will v|be adv:adj|eventual-LY v|be adj|developed .

%snd: <04.24.13><04.33.90>

%JCC: and as a result , (0.63) er a mutation of the liver cell will occur errfr ::: and cancer (0.60) will {be} ~ eventually be developed . errfr ::: |

*JCC: however , not all of the consequences of virus infection mmm are that severe .

%mor: adv:wh|however neg|not qn|all prep|of det|the n|consequence-PL prep|of n|virus n|infection fil|mmm v|be&PRES det|that adj|severe .

%snd: <04.35.65><04.42.72>

%JCC: however , not all of the consequences of virus infection (0.62) mmm are that severe . errfr ::: |

*JCC: for example, chicken pox, which is caused by the varicella virus .

%mor: prep|for n|example n|chicken n|pox rel|which aux|be&3S

part|cause-PERF prep|by det|the ?|varicella n|virus .

%snd: <04.43.31><04.49.27>

%JCC: for example , chicken pox errfr ::: (0.57) which is caused by the varicella virus . errfr :;:a |

*JCC: .

%mor: fil|mmm ?|hu# fil|mmm .

%snd: <04.50.02><04.53.32>

%JCC: mmm {hu@} # mmm . |

*JCC: yes it is a self-limiting disease in that infant or children infected with that particular virus will heal itself .

%mor: co|yes pro|it v|be&3S det|a ?|self-limiting dis#n|ease^n|disease prep|in pro:dem|that adj|infant conj:coo|or n|child&PL v|infect-PAST prep|with pro:dem|that adj|particular n|virus aux|will v|heal pro:refl|itself .

%snd: <04.55.11><05.08.39>

%JCC: yes mmm (0.45) it is a self-limiting disease errfr ::: in that er (0.68) infant {or} * (0.87) or children infected with that particular virus err_m_m :;;m will heal itself . err_m_r :;:a |

*JCC: and after which the disease will rarely develop again .

%mor: conj:coo|and prep|after rel|which det|the dis#v|ease aux|will adv:adj|rare-LY v|develop adv|again .

%snd: <05.09.76><05.15.75>

%JCC: and (0.56) after which er the disease will rarely develop again . errfr ::: |

*JCC: then to general procedure .

%mor: adv:tem|then prep|to n|general n|procedure .

%snd: <05.17.71><05.21.04>

%JCC: mmm (0.82) then to general procedure . err_m_s ::: |

*JCC: there are several ways to treat virus infection in human body .

%mor: pro:exist|there v|be&PRES qn|several n|way-PL prep|to n|treat n|virus n|infection prep|in n|human n|body .

%snd: <05.21.82><05.27.66>

%JCC: mmm (0.85) there are several ways errfr ::: to treat virus infection in human body . errfr :;;a |

*JCC: the most direct way is using anti-viral drugs which act in different way .

%mor: det|the qn|most adj|direct n|way v|be&3S part|use-PROG adj|anti-viral n|drug-PL det:wh|which v|act prep|in adj|different n|way .

%snd: <05.28.56><05.37.44>

%JCC: er the most direct way is using anti-viral drugs errfr ::: (0.79) er which act in (0.57) different way . err_m_m :;:a |

*JCC: .

%mor: fil|mmm .

%snd: <05.38.26><05.38.82>

%JCC: mmm . |

*JCC: one type of antiviral drugs prevents but stops the attachment stage of the viral infection process whereas some other drugs stop various replication stage of the viral infection process in the human body .

%mor: det:num|one n|type prep|of adj|antiviral n|drug-PL n|vent-PL conj:coo|but n|stop-PL det|the n|attachment n|stage prep|of det|the ?|viral n|infection n|process conj:subor|whereas qn|some qn|other n|drug-PL n|stop adj|various n|replication n|stage prep|of det|the ?|viral n|infection n|process prep|in det|the n|human n|body .

%snd: <05.40.09><06.00.76>

%JCC: one type of antiviral drugs er (0.54) prevents er but stops (0.45) er the attachment stage of the viral infection process err_m_l ::: (0.73) whereas some other drugs stop various replication stage of the viral infection process in the human body . err_m_m ::: |

*JCC: for example some stop their RNA replication stage whereas some others stop the DNA replication stage .

%mor: prep|for n|example qn|some v|stop pro:poss:det|their n|RNA n|replication n|stage conj:subor|whereas qn|some pro:indef|other-PL v|stop det|the n|DNA n|replication n|stage .

%snd: <06.01.42><06.09.04>

%JCC: for example some stop their RNA replication stage errfr ::: whereas some others stop (0.44) the DNA replication stage . errfr ::: |

*JCC: .

%mor: fil|mmm .

%snd: <06.10.19><06.10.77>

%JCC: mmm . |

*JCC: however, antiviral drugs are often very expensive .

%mor: adv|however ?|antiviral n|drug-PL v|be&PRES adv|often adv:int|very adj|expensive .

%snd: <06.11.41><06.16.39>

%JCC: however (0.53) antiviral drugs are often very expensive . errfr ::: |

*JCC: and there are very few in the market .

%mor: conj:coo|and adv:loc|there v|be&PRES adv:int|very qn|few adv:loc|in det|the n|market .

%snd: <06.16.99><06.22.82>

%JCC: {and there are not} ~ (0.91) er and there are very few in the market .

errfr ::: |

*JCC: therefore, most of the time, symptomatic treatment is used .

%mor: adv|therefore qn|most prep|of det|the n|time adj|symptomatic n|treatment v|be&3S adv|used .

%snd: <06.23.38><06.27.78>

%JCC: therefore, most of the time , symptomatic treatment is used . errfr ::: |

*JCC: that is, the patient is treated according to their symptoms .

%mor: pro:dem|that v|be&3S det|the n|patient aux|be&3S part|treat-PERF part|accord-PROG prep|to pro:poss:det|their n|symptom-PL .

%snd: <06.28.65><06.34.01>

%JCC: that is errfr :::; er the patient is treated according (0.52) to their symptoms . errfr ::: |

*JCC: for examples, if a patient er develop a fever, they will be dehydrated, so they will be given drugs to combat that condition .

%mor: prep|for n|example-PL conj:subor|if det|a n|patient fil|er v|develop det|a n|fever pro|they aux|will v|be part|hydrate-PERF co|so pro|they aux|will v|be part|give&PERF v|drug-3S prep|to n|combat det|that n|condition .

%snd: <06.34.85><06.49.71>

%JCC: for examples , if a patient (0.76) er develop a fever , errfr :;:b er they will be dehydrated , errfr ::: (0.60) so er they will be given er (0.60) drugs errfr ::: to combat that condition . errfr :;;a |

*JCC: in another case , if a patient is infected with influenza, then they might develop a soar throat, in which case cough syrup or other soothing solution are given .

%mor: prep|in det|another n|case conj:subor|if det|a n|patient aux|be&3S part|infect-PERF prep|with n|influenza adv:tem|then pro|they aux|might v|develop det|a v|soar n|throat prep|in rel|which v|case v|cough n|syrup conj:coo|or qn|other part|soothe-PROG n|solution aux|be&PRES part|give&PERF .

%snd: <06.50.55><07.12.79>

%JCC: er in another case er , if a patient is infected with influenza , errfr :;:b (0.60) then (0.95) they might develop a soar throat , errfr ::: (0.47) in which case er (1.03) {cough syr@} * cough syrup (0.46) or other soothing solution are given . err_m_s :;:a |

@End

Sample 5 (Computer major, unplanned, unfamiliar)

*OCT: for infection of virus in a human body, the most common cause is the person having a weak body .

%mor: prep|for n|infection prep|of n|virus prep|in det|a n|human n|body det|the qn|most adj|common n|cause v|be&3S det|the n|person part|have-PROG det|a adj|weak n|body .

%snd: <00.23.97><00.40.05>

%OCT: mmm for infection of virus in a human body (0.94) mmm the most (0.63) {common} * (0.52) common cause {is (0.57) having} * is err_m_s ::: the person having a weak body . errfr :;;a |

*OCT: that is he or she doesn't have enough antibodies in their body .

%mor: pro:dem|that v|be&3S pro|he conj:coo|or pro|she v|doesn't v|have adj|enough n|antibody-PL^n|body-PL prep|in pro:poss:det|their n|body .

%snd: <00.40.15><00.51.24>

%OCT: that is errfr :::; (0.82) {he} * (0.66) he or she (0.62) doesn't have enough (0.85) antibodies in their body er . errfr ::: |

*OCT: not having enough rest .

%mor: neg|not part|have-PROG adj|enough n|rest .

%snd: <00.52.28><00.55.47>

%OCT: not having enough rest {and} # . errfr :;;a |

*OCT: that is the weak body .

%mor: pro:dem|that v|be&3S det|the adj|weak n|body .

%snd: <00.56.20><01.02.66>

%OCT: that is (2.86) the (1.18) weak body . errfr ::: |

*OCT: weak body then virus can get .

%mor: adj|weak n|body adv:tem|then n|virus aux|can v|get .

%snd: <01.03.24><01.08.54>

%OCT: and weak body then virus can get # . err_a err_m_s ::: |

*OCT: in fact virus and bacteria are everywhere in our life .

%mor: prep|in n|fact n|virus conj:coo|and n|bacteria v|be&PRES n|everywhere prep|in pro:poss:det|our n|life .

%snd: <01.10.42><01.18.11>

%OCT: in fact {virus are} ~ (1.12) virus and bacteria are everywhere (0.82) in our life . err_m_m ::: |

*OCT: for a weak body those virus and bacterias can get into our skin .

%mor: prep|for det|a adj|weak n|body det|those n|virus conj:coo|and n|bacteria-PL aux|can v|get prep|into pro:poss:det|our n|skin .

%snd: <01.19.03><01.30.50>

%OCT: er for a weak body (0.63) those virus and bacterias (0.52) can (0.55) get (0.57) into (0.51) er {our} * our skin . err_m_m err_m_r ::: |

*OCT: and then our skin through our mouth, through our eyes and nose and etc .

%mor: conj:coo|and adv:tem|then pro:poss:det|our n|skin prep|through pro:poss:det|our n|mouth prep|through pro:poss:det|our n|eye-PL conj:coo|and v|nose conj:coo|and adv|etc .

%snd: <01.31.47><01.49.11>

%OCT:and then (0.64) {we may} * (1.32) {we may get those} # (2.01) {our skin} * (0.80) our skin through our mouth (1.30) through our eyes and nose and etc . err_a_s err_s_r ::: |

*OCT: then those virus will maybe attack our body cells .

%mor: adv:tem|then det|those n|virus v|will adv|maybe v|attack pro:poss:det|our n|body n|cell-PL .

%snd: <01.49.82><02.07.23>

%OCT: then (0.95) those virus will$ (0.81) mmm (0.74) maybe {attach on (0.60) our} ~ attack {our} * (1.12) our {body shells} ~ body cells . err_m_m ::: |

*OCT: if our anti-virus is stop in our body, is not powerful at that time, then it will be easily get infected by even a very weak virus .

%mor: conj:subor|if pro:poss:det|our ?|anti-virus v|be&3S n|stop prep|in pro:poss:det|our n|body v|be&3S neg|not adj|powerful prep|at det|that n|time adv:tem|then pro|it aux|will v|be adv:adj|easy-LY v|get part|infect-PERF prep|by adj|even det|a adv:int|very adj|weak n|virus .

%snd: <02.08.02><02.39.09>

%OCT: and$ (0.56) {our} ~ if our {anti@} * (1.19) {anti-virus} * {anti-virus} * (1.41) anti-virus (1.85) is stop in our body err_m_l err_m_s :;:b is {not} * (0.91) not powerful at that time err_m_s :;:b then (0.69) it will be easily get infected {by those} ~ (0.97) {even} ~ by even a very weak virus . err_m_s ::: |

*OCT: the possible consequences are you will most probably not have bad health condition .

%mor: det|the adj|possible n|consequence-PL v|be&PRES adv|maybe pro|you v|will qn|most adv:adj|probable-LY neg|not v|have adj|bad n|health n|condition .

%snd: <02.40.35><03.13.04>

%OCT: and the possible consequences (1.04) are errfr ::: (1.09) {maybe you will have} ~ (2.38) you will (0.76) most probably {have} ~ (1.98) er (2.68) mmm$ (1.06) not (1.25) have (2.78) bad {health condition} * health condition . err_m_r :;:a |

*OCT: you will infected then you will get ill .

%mor: pro|you v|will part|infect-PERF adv:tem|then pro|you aux|will v|get adj|ill .

%snd: <03.14.20><03.21.85>

%OCT: you will infected err_m_s ::: then you will {get illt} * (0.95) {get ill} * (1.30) get ill . errfr ::: |

*OCT: if you don't have enough rest or consult doctor as soon as possible , then the virus will weaken your anti-virus system in your body .

%mor: conj:subor|if pro|you v|don't v|have adj|enough n|rest conj:coo|or v|consult n|doctor prep|as adv|soon prep|as adj|possible adv:tem|then det|the n|virus aux|will v|weaken pro:poss:det|your ?|anti-virus n|system prep|in pro:poss:det|your n|body .

%snd: <03.21.95><03.41.39>

%OCT: if {the} # you don't have enough rest errfr :;:b or consult doctor as soon as possible err_m_s :;:b then (0.65) the virus will weaken your (0.49) {anti-virus sy@} * (0.78) anti-virus system {in your body} * (0.72) in your (0.67) body . errfr ::: |

*OCT: then more and more virus or bacteria will infect you .

%mor: adv:tem|then qn|more conj:coo|and qn|more n|virus conj:coo|or n|bacteria aux|will v|infect pro|you .

%snd: <03.41.40><03.49.30>

%OCT: then (0.54) er (0.46) more and more virus {and} rpl or bacteria will infect you . err_m_m ::: |

*OCT: then finally you will get many disease and illnesses at the same time .

%mor: adv:tem|then adv:adj|final-LY pro|you aux|will v|get qn|many n|disease conj:coo|and n:adj|ill-NESS-PL prep|at det|the adj|same n|time .

%snd: <03.50.07><03.58.25>

%OCT: then (0.95) mmm finally you will get many disease and illnesses at the same time . err_m_m ::: |

*OCT: then your body cannot cope with these .

%mor: adv:tem|then pro:poss:det|your n|body aux|can~neg|not v|cope prep|with pro:dem|these .

%snd: <03.58.27><04.05.13>

%OCT: then your body {will not} ~ (0.77) cannot {cope with} * (1.02) cope with (1.10) these # . err_a_l |

*OCT: the ultimate result maybe is death .

%mor: adj|ultimate n|result adv|maybe v|be&3S n|death .

%snd: <04.06.42><04.12.57>

%OCT: the {ultimate (1.08) go@} ~ (0.46) ultime result maybe is death . errfr ::: |

*OCT: for general procedure of dealing of a virus infected person .

%mor: prep|for n|general n|procedure prep|of part|deal-PROG prep|of det|a n|virus v|infect-PAST n|person .

%snd: <04.13.86><04.21.01>

%OCT: mmm for general procedure of dealing of a virus infected person . err_m_s ::: |

*OCT: for most proper way is to consult doctor when you see any problems in .

%mor: prep|for qn|most adj|proper n|way v|be&3S inf|to v|consult n|doctor conj:subor|when pro|you v|see qn|any n|problem-PL adv:loc|in .

%snd: <04.21.92><04.40.86>

%OCT: mmm {for} # (2.45) {most} * (2.05) most proper way is err_m_s ::: to consult doctor err_m_m :;;a when you see {any} * (0.80) {any$} (1.32) any problems in # . err_a_l :;:a |

*OCT: you find it .

%mor: pro|you v|find pro|it .

%snd: <04.41.55><04.44.28>

%OCT: you find it errfr ::: {you can} # . |

*OCT: listen to the doctor's advice .

%mor: v|listen prep|to det|the n|doctor's n|advice .

%snd: <04.47.87><04.54.92>

%OCT: er$ {if the} # (1.42) listen to the doctor's advice . errfr ::: |

*OCT: this virus is highly infected, then you may have to wear masks or to be guarantee .

%mor: det|this n|virus v|be&3S adv:adj|high-LY part|infect-PERF adv:tem|then pro|you aux|may v|have inf|to v|wear n|mask-PL conj:coo|or inf|to v|be n|guarantee .

%snd: <04.56.07><05.16.23>

%OCT: {if the doctor ask you never} # (0.85) {the vi@} ~ this virus (0.71) is (0.49) highly infected err_m_s err_m_r ::: then {you may} * (0.79) you may have to wear mases err_m_m ::: {or} * or to be {gua@} * {guarantee} * guarant@ err_s_l err_m_s :;;a |

*OCT: to be keep away from all your family or friends .

%mor: ,inf|to v|be v|keep adv|away prep|from qn|all pro:poss:det|your n|family conj:coo|or n|friend-PL .

%snd: <05.17.30><05.25.73>

%OCT: er to be keep away from all your family {or} * (1.14) or friends . err_m_m :;;a |

*OCT: .

%mor: .

%snd: <05.26.80><05.32.72>

%OCT: mmm (0.77) {for dealing} # {then} # (0.77) {dealing with} # . |

*OCT: if doctors give you medicines or those treatment, you must follow them strictly .

%mor: conj:subor|if v|doctor-3S n|give pro|you n|medicine-PL conj:coo|or det|those n|treatment pro|you aux|must v|follow pro|them adv:adj|strict-LY .

%snd: <05.34.42><05.56.80>

%OCT: mmm {if a doctor} ~ if doctors give you (0.90) er medicines {or} * (1 22) {or} * or those treatment err_m_m :;:b you must er {follow them} * (1.06) follow them$ {strictly} * er (0.99) strictly . errfr ::: |

*OCT: if you do not obbey them, the virus maybe get antibody of those treatments .

%mor: conj:subor|if pro|you aux|do neg|not v|abbey pro|them det|the n|virus adv|maybe v|get n|antibody^n|body prep|of det|those n|treatment-PL .

%snd: <05.58.25><06.31.27>

%OCT: if you do not obbey them errfr :;:b the virus {may get} ~ (1.33) {maybe get anti@ * } ~ (1.10) {anti@} * (3.84) {maybe get} * (4.75) {maybe (2.74) be get antibody (0.70) of} * maybe get antibody of those treatments . err_s_l ::: |

*OCT: then the virus is stronger and stronger .

%mor: adv:tem|then det|the n|virus v|be&3S adj|strong-CP conj:coo|and adj|strong-CP .

%snd: <06.31.44><06.35.10>

%OCT: then the virus is (0.68) stronger and stronger . errfr ::: |

*OCT: they may mutate you, then the virus will be more hard to cope with and manipulate .

%mor: pro|they aux|may v|mutate pro|you adv:tem|then det|the n|virus aux|will v|be qn|more adj|hard inf|to v|cope prep|with conj:coo|and v|manipulate .

%snd: <06.36.06><06.46.83 >

%OCT: and they may mutate you err_s_r ::: then the virus will be more {hard to} * (0.60) hard err_m_m ::: to (1.13) cope with and manipulate . errfr :;;a |

*OCT: so the prevention of getting infected by virus is have good rest .

%mor: adv|so det|the n|prevention prep|of part|get-PROG part|infect-PERF prep|by n|virus v|be&3S v|have adj|good n|rest .

%snd: <06.47.95><07.00.46>

%OCT: so (1.05) the prevention {of} * (1.50) of (0.81) getting infected by virus errfr :;;b is er (1.36) have good rest . err_m_s ::: |

*OCT: have proper diet .

%mor: v|have adj|proper n|diet .

%snd: <07.00.97><07.05.42>

%OCT: have er (1.61) proper diet . errfr ::: |

*OCT: .

%mor: .

%snd: <07.06.77><07.08.78>

%OCT: {never} * .

*OCT: and do exercise frequently in order to have healthy body to fight against those virus .

%mor: conj:coo|and v|do n|exercise adv:adj|frequent-LY prep|in n|order inf|to v|have adj|healthy n|body prep|to n|fight prep|against det|those n|virus .

%snd: <07.10.89><07.24.04>

%OCT: {ne@} # and do exercise frequently errfr ::: (0.45) {to have} ~ in order to have healthy body err_m_m :;;a (0.93) to fight (0.54) against those virus . err_m_m :;;a |

@End

Sample 6 (Computer major, unplanned, familiar)

*JYL: the virus can enter the body through three main routes .

%mor: det|the n|virus aux|can v|enter det|the n|body prep|through det:num|three adj|main n|route-PL .

%snd: <00.15.16><00.21.73>

%JYL: the virus can enter the body mmm through (1.07) er three main routes . errfr ::: |

*JYL: and they may use some agents to go into our bodies, for example, air, food and water and body fluid .

%mor: conj:coo|and pro|they aux|may v|use qn|some n|agent-PL inf|to v|go prep|into pro:poss:det|our n|body-PL prep|for n|example n|air n|food conj:coo|and v|water conj:coo|and n|body n|fluid .

%snd: <00.22.23><00.33.54>

%JYL: and they may mmm (0.47) er (0.51) use some agents errfr ::: to go into our bodies , for example , air , er food and (0.43) water and body fluid . err_m_s :;;a |

*JYL: cold and influenza are the examples of how virus go to our body through the air .

%mor: n|cold conj:coo|and n|influenza v|be&PRES det|the n|example-PL prep|of adv:wh|how n|virus v|go prep|to pro:poss:det|our n|body prep|through det|the n|air .

%snd: <00.34.12><00.43.08>

%JYL: mmm cold and influenza are the examples of err_m_m ::: (0.71) er how virus er (1.02) go to our body through the air . err_m_m :;:a |

*JYL: when we cough we may spread the saliva and virus .

%mor: conj:subor|when pro|we v|cough pro|we aux|may n|spread det|the n|saliva conj:coo|and n|virus .

%snd: <00.43.60><00.47.57>

%JYL: mmm when we cough errfr :;:b we may spread the saliva and virus . errfr ::: |

*JYL: .

%mor: .

%snd: <00.48.05><00.48.51>

%JYL: and_pfp . |

*JYL: it may go into others people's body and go into the respiratory system of us .

%mor: pro|it aux|may v|go prep|into pro:indef|other-PL v|people~v|be&3S n|body conj:coo|and v|go prep|into det|the adj|respiratory n|system prep|of pro|us .

%snd: <00.49.87><00.58.40>

%JYL: it may go into er others people's body err_m_m err_m_m ::: {and} * (0.64) and go into the respiratory system of us . err_m_l ::: |

*JYL: also the wind from the northern China may blow to Hong Kong and then it may carry some virus .

%mor: adv|also det|the n|wind prep|from det|the adj|northern n:prop|China aux|may v|blow prep|to n:prop|Hong n:prop|Kong conj:coo|and adv:tem|then pro|it aux|may v|carry qn|some n|virus .

%snd: <00.58.92><01.07.02>

%JYL: also the wind from the {nor@} * {northern} * northern China may blow to Hong Kong err_m_r ::: and then it may carry some virus . err_m_m ::: |

*JYL: also the other route is through the food and water .

%mor: adv|also det|the qn|other n|route v|be&3S prep|through det|the n|food conj:coo|and v|water .

%snd: <01.08.63><01.14.11>

%JYL: also mmm the other route is (0.43) through the (0.40) mmm food and water . err_m_r ::: |

*JYL: for example seafood sushi, they are all high-risk food .

%mor: prep|for n|example n|seafood n|sushi pro|they aux|be&PRES qn|all adj|high-risk n|food .

%snd: <01.14.35><01.19.94>

%JYL: for example (0.42) seafood er sushi , they are all high-risk (0.42) mmm food . err_m_d ::: |

*JYL: and they may en# the virus may enter our digestive system .

%mor: conj:coo|and pro|they aux|may ?|en# det|the n|virus aux|may v|enter pro:poss:det|our adj|digestive n|system .

%snd: <01.20.21><01.25.02>

%JYL: and {they may en@} rpl the virus may enter our er (0.47) digestive system . errfr ::: |

*JYL: for the last one it is through the body fluid .

%mor: prep|for det|the adj|last pro:indef|one pro|it v|be&3S prep|through det|the n|body n|fluid .

%snd: <01.26.25><01.30.80>

%JYL: mmm {for} * (0.56) for the last one it is through the body fluid . err_m_s ::: |

*JYL: and we may through some sex and then we may spread the virus .

%mor: conj:coo|and pro|we aux|may prep|through qn|some n|sex conj:coo|and adv:tem|then pro|we aux|may n|spread det|the n|virus .

%snd: <01.31.36><01.39.79>

%JYL: and {we may use the} ~ (0.44) er we may through some (0.59) sex err_m_s ::: and$ then we may spread the virus . errfr ::: |

*JYL: and it may directly enter the circulatory system .

%mor: conj:coo|and pro|it aux|may adv:adj|direct-LY v|enter det|the adj|circulatory n|system .

%snd: <01.40.17><01.44.41>

%JYL: and it may directly enter the er circulatory system . errfr ::: |

*JYL: and the virus can go into our circulatory system and affect our body finally .

%mor: conj:coo|and det|the n|virus aux|can v|go prep|into pro:poss:det|our adj|circulatory n|system conj:coo|and v|affect pro:poss:det|our n|body adv:adj|final-LY .

%snd: <01.44.88><01.54.49>

%JYL: and {all} * {all three routes may} # (0.53) er the virus can go into our circulatory system errfr ::: and affect our body (0.84) finally . errfr ::: |

*JYL: and for how we treat with them is that if the disease is easy spread, we may isolate them and ask them to wear mask and stay at home .

%mor: conj:coo|and prep|for adv:wh|how pro|we v|treat prep|with pro|them v|be&3S pro:dem|that conj:subor|if det|the dis#n|ease^n|disease v|be&3S adj|easy n|spread pro|we aux|may v|isolate pro|them conj:coo|and v|ask pro|them prep|to n|wear n|mask conj:coo|and v|stay prep|at n|home .

%snd: <01.55.35><02.09.01>

%JYL: and for how we treat with them err_m_s :;:m is that er (0.62) if th[illegible]sease is (0.62) easy spread err_m_m :;:b (0.65) we may isolate them errfr ::: and ask them errfr ::: to wear mask err_m_s :;;a and stay at home . errfr :;;a |

*JYL: and how we treat them .

%mor: conj:coo|and adv:wh|how pro|we v|treat pro|them .

%snd: <02.09.47><02.11.94>

%JYL: and how we er treat them . err_m_s ::: |

*JYL: we may ask them to see the doctors .

%mor: pro|we aux|may v|ask pro|them inf|to v|see det|the n|doctor-PL .

%snd: <02.12.32><02.17.72>

%JYL: er {we may take} rpl (0.51) we may ask them errfr ::: to (0.52) see the doctors {and} * . errfr :;;a |

*JYL: and tell them to take more rest and eat some healthy food and injection or something like that .

%mor: conj:coo|and v|tell pro|them inf|to v|take qn|more n|rest conj:coo|and v|eat qn|some adj|healthy n|food conj:coo|and n|injection conj:coo|or pro:indef|something prep|like pro:dem|that .

%snd: <02.18.27><02.26.77>

%JYL: and tell them errfr ::: to take more rest errfr :;;a and eat some healthy food and (0.77) injection or something like that . err_m_s :;;a |

*JYL: It may help them to get better health .

%mor: n:prop|It aux|may v|help pro|them inf|to v|get adv|better n|health .

%snd: <02.27.28><02.31.40>

%JYL: It may help them errfr ::: to (0.99) er get better health . errfr :;;a |

@End

Sample 7 (Computer major, planned, unfamiliar)

*LSZ: the general process of the infection of the virus in a human body can be by air, though some virus cannot survive in the air longer than a few seconds .

%mor: det|the n|general n|process prep|of det|the n|infection prep|of det|the n|virus prep|in det|a n|human n|body aux|can v|be prep|by n|air adv|though qn|some n|virus aux|can~neg|not v|survive prep|in det|the n|air adj|long-CP prep|than det|a qn|few n|second-PL .

%snd: <00.14.02><00.22.12>

%LSZ: er the general process of the infection of the virus in a human body can be by air err_m_s ::: though some virus cannot survive in the air longer than a few seconds . err_m_m :;:a |

*LSZ: or by contact with patient of the source of viruses .

%mor: conj:coo|or prep|by n|contact prep|with n|patient prep|of det|the n|source prep|of n|virus-PL .

%snd: <00.22.58><00.25.72>

%LSZ: or by contact with patient of the source of viruses . err_m_s ::: |

*LSZ: the consequences of infecting the viruses can be .

%mor: det|the n|consequence-PL prep|of part|infect-PROG det|the n|virus-PL aux|can v|be .

%snd: <00.26.85><00.29.88>

%LSZ: er of infecting the viruses errfr :;;m the consequences can be . err_m_s :::

|

*LSZ: the virus will attack your body .

%mor: det|the n|virus aux|will v|attack pro:poss:det|your n|body .

%snd: <00.30.21><00.32.19>

%LSZ: er the virus will attack your body . errfr ::: |

*LSZ: and then the immune system will defense .

%mor: conj:coo|and adv:tem|then det|the adj|immune n|system v|will n|defense .

%snd: <00.32.62><00.34.55>

%LSZ: and then the immune system will defense . errfr ::: |

*LSZ: the immune system will produce some white blood cell to engulf the virus to destroy it .

%mor: det|the adj|immune n|system aux|will v|produce qn|some adj|white n|blood n|cell inf|to v|engulf det|the n|virus inf|to v|destroy pro|it .

%snd: <00.35.69><00.41.26>

%LSZ: {they} rpl er the immune system will produce some white blood cell err_m_m ::: to engulf the virus errfr :;;a to destroy it . errfr :;;a |

*LSZ: and your body would have some syndromes such as cough, high body temperature, and maybe headache .

%mor: conj:coo|and pro:poss:det|your n|body aux|will&COND v|have qn|some n|syndrome-PL qn|such prep|as n|cough adj|high n|body n|temperature conj:coo|and adv|maybe n|headache .

%snd: <00.41.60><00.47.50>

%LSZ: and your body {will} ~ would have some syndromes (0.51) such as cough , high body temperature , and maybe headache . err_m_l ::: |

*LSZ: in some serious cases with the strong virus, the immune system fail to defense .

%mor: prep|in qn|some adj|serious n|case-PL prep|with det|the adj|strong n|virus det|the adj|immune n|system v|fail prep|to n|defense .

%snd: <00.48.27><00.52.83>

%LSZ: in some serious cases with the strong virus (0.49) the immune system fail to defense . err_m_s ::: |

*LSZ: and the viruses will attack your organs which may fail to work .

%mor: conj:coo|and det|the n|virus-PL v|will v|attack pro:poss:det|your n|organ-PL rel|which aux|may v|fail inf|to v|work .

%snd: <00.53.12><00.56.41>

%LSZ: and the viruses will attack your organs errfr ::: which may fail to work . errfr :;:a |

*LSZ: if not cure the patient in time, the patient may die due to the failure of working of organs .

%mor: conj:subor|if neg|not v|cure det|the n|patient prep|in n|time det|the n|patient aux|may n|die adj|due infl|to det|the n|failure prep|of part|work-PROG prep|of n|organ-PL .

%snd: <00.56.88><01.03.37>

%LSZ: if not cure the patient in time err_m_s :;:b the patient may die errfr ::: due to the failure of working of organs . errfr :;;a |

*LSZ: the procedures to deal with the virus infected person can use the anti-virus inject to patient's body .

%mor: det|the n|procedure-PL prep|to n|deal prep|with det|the n|virus v|infect-PAST n|person aux|can v|use det|the ?|anti-virus v|inject prep|to n|patient~v|be&3S n|body .

%snd: <01.04.29><01.12.24>

%LSZ: {and how} # the procedures to deal with the virus infected person errfr :;;b can use the anti-virus inject {to the body} ~ to patient's body . err_m_s err_m_l ::: |

*LSZ: and the blood anti-viruses add at the the immune system to defense .

%mor: conj:coo|and det|the n|blood ?|anti-viruses v|add adv:loc|at det|the det|the adj|immune n|system prep|to n|defense .

%snd: <01.12.49><01.16.98>

%LSZ: and the blood anti-viruses add at the the immune system to defense . err_s_s err_s_l ::: |

@End

Sample 8 (Computer major, planned, familiar)

*CYN: ok basically the infection of virus in computer is come from Internet .

%mor: coo|ok adv:adj|basic-LY det|the n|infection prep|of n|virus prep|in n|computer v|be&3S v|come prep|from n:prop|Internet .

%snd: <00.17.80><00.24.46>

%CYN: ok er basically the er infection of virus in computer is come from er Internet . err_m_s ::: |

*CYN: why I said Internet because Internet is just the windows to allows you to touch the outside of the world .

%mor: adv:wh|why pro|I v|say&PAST n:prop|Internet conj:subor|because n:prop|Internet v|be&3S adv:int|just det|the n|window-PL prep|to v|allow-3S pro|you inf|to v|touch det|the out#n|side prep|of det|the n|world .

%snd: <00.24.67><00.32.91>

%CYN: {why} * why I said Internet err_m_s ::: because er Internet is just the windows err_m_s err_m_m :;:a (0.43) er {to} * to allows you err_m_s :;;a to touch the outside of {the} * the world . errfr :;;a |

*CYN: and that means you have an interaction between the outside of the world and your computer itself .

%mor: conj:coo|and pro:dem|that v|mean-3S pro|you v|have det|a n|interaction prep|between det|the out#n|side prep|of det|the n|world conj:coo|and pro:poss:det|your n|computer pro:refl|itself .

%snd: <00.33.42><00.38.27>

%CYN: and that means errfr ::: you have an interaction between the outside of the world and {you} rpl your computer itself . errfr :;:a |

*CYN: so they just give other a chance to give to transfer the virus into your computers .

%mor: co|so pro|they adv:int|just v|give qn|other det|a n|chance prep|to n|give inf|to v|transfer det|the n|virus prep|into pro:poss:det|your n|computer-PL .

%snd: <00.39.62><00.48.56>

%CYN: so they just {gave you a * a chance} rpl give other a chance err_m_m ::: {to} * {to give} rpl to transfer the virus {to} rpl {in your} rpl into your computers . errfr :;;a |

*CYN: so they make a big problems .

%mor: co|so pro|they v|make det|a adj|big n|problem-PL .

%snd: <00.48.88><00.50.59>

%CYN: so they make a big problems . err_m_m ::: |

*CYN: so from Internet in nowadays we also basically have three or four kinds of activities like we always access some websites .

%mor: adv|so prep|from n:prop|Internet adv:loc|in adv:tem|nowadays pro|we adv|also adv:adj|basic-LY v|have det:num|three conj:coo|or det:num|four n|kind-PL prep|of n|activity-PL prep|like pro|we adv|always n|access qn|some n|websites .

%snd: <00.50.96><01.00.91>

%CYN: so er from Internet er in nowadays we also basically have$ three or four kinds of activities err_m_s ::: like we always access some websites . err_m_l :;:a |

*CYN: or we receive or send emails .

%mor: conj:coo|or pro|we v|receive conj:coo|or v|send n|email-PL .

%snd: <01.01.50><01.04.72>

%CYN: er {or} * or we receive or send emails . errfr ::: |

*CYN: and use instant communication software to communicate with others like MSN messengers or ICQ .

%mor: conj:coo|and v|use n|instant n|communication n|software inf|to v|communicate prep|with pro:indef|other-PL v|like n:prop|MSN n|messenger-PL conj:coo|or n:prop|ICQ .

%snd: <01.05.10><01.12.92>

%CYN: and use instant communication software errfr ::: to communicate with others like MSN messengers or ICQ . errfr :;;a |

*CYN: and the last thing should be downloading some files or MP3 or MP4 , MV so on .

%mor: conj:coo|and det|the adj|last n|thing aux|should v|be part|download-PROG qn|some n|file-PL conj:coo|or n:prop|MP3 conj:coo|or n:prop|MP4 n:prop|MV adv|so adv:loc|on .

%snd: <01.13.41><01.23.52>

%CYN: and$ er the last thing {should be} * should be downloading some files or MP3 or er er MP4 , MV so so on {because} # er # . err_a_d ::: |

*CYN: so let's say when you access website like Yahoo or somewhere , you have to download the picture and words or something .

%mor: conj:subor|so v|let~pro|us v|say conj:subor|when pro|you n|access n|website v|like n:prop|Yahoo conj:coo|or n|somewhere pro|you v|have inf|to v|download det|the n|picture conj:coo|and n|word-PL conj:coo|or pro:indef|something .

%snd: <01.24.04><01.36.76>

%CYN: so er let's say errfr ::: er {when we access the website} rpl when you access website like Yahoo or somewhere err_m_m :;:b you have to er download the er (0.42) picture and words or something . errf_m_m ::: |

*CYN: so you can see the website , you can read it .

%mor: co|so pro|you aux|can v|see det|the ?|website pro|you aux|can re#n|ad pro|it .

%snd: <01.37.14><01:41.03>

%CYN: er so you can see the website , errfr ::: {you can} * you can{ read} * read it or ok_pfp . errfr ::: |

*CYN: so in that process you will through downloading some file , that means .

%mor: adv|so prep|in det|that n|process pro|you v|will adv:loc|through part|download-PROG qn|some n|file pro:dem|that v|mean-3S .

%snd: <01.41.49><01.46.05>

%CYN: so in that process you will er through downloading some file err_m_s ::: that means . errfr ::; |

*CYN: and if the file contain some virus , so your computer will get infections .

%mor: conj:coo|and conj:subor|if det|the n|file v|contain qn|some n|virus adv|so pro:poss:det|your n|computer aux|will v|get n|infection-PL .

%snd: <01.46.08><01.52.02>

%CYN: and if the file contain some (0.41) virus err_m_s :;:b so {your} * your computer will get infections . errfr ::: |

*CYN: and emails is simply XX .

%mor: conj:coo|and n|email-PL v|be&3S adv:adj|simple-LY adj|XX .

%snd: <01.51.63><01.55.72>

%CYN: and emails {is} * er is simply XX . err_m_s err_i_l ::: |

*CYN: I think everyone get a lot of email every day .

%mor: pro|I v|think pro:indef|everyone v|get det|a n|lot prep|of n|email qn|every n|day .

%snd: <01.55.74><01.58.59>

%CYN: I think errfr :::; everyone get a lot of email every day . err_m_s err_m_m ::: |

*CYN: and some of the emails may be unknown emails .

%mor: conj:coo|and qn|some prep|of det|the n|email-PL aux|may v|be n|unknown n|email-PL .

%snd: <01.58.74><02.01.60>

%CYN: and some of the emails may be unknown emails . errfr ::: |

*CYN: that mean you don't know who send it to you .

%mor: pro:dem|that v|mean pro|you aux|do~neg|not v|know pro:wh|who v|send pro|it prep|to pro|you .

%snd: <02.01.76><02.04.08>

%CYN: that mean err_m_s ::: you don't know errfr :;:a {who} * who send it to you . err_m_s :;:a |

*CYN: and the topic you are don't knows .

%mor: conj:coo|and det|the n|topic pro|you aux|be&PRES aux|do~neg|not v|know-3S .

%snd: <02.04.82><02.07.01>

%CYN: and the topic {you} * er you are don't knows {and} # . err_m_s ::: |

*CYN: so that have some probabilities of virus get inside it .

%mor: conj:subor|so rel|that v|have qn|some n|probability-PL prep|of n|virus aux|get prep|inside pro|it ..

%snd: <02.07.77><02.12.78>

%CYN: so that have some probabilities err_m_s ::: {in} ~ of virus get inside it . err_m_s :;;a |

*CYN: so if you open the emails you will just download the virus in your computer , so you will get infection of your computers .

%mor: co|so conj:subor|if pro|you v|open det|the n|email-PL pro|you v|will adv:int|just v|download det|the n|virus prep|in pro:poss:det|your n|computer conj:subor|so pro|you aux|will v|get n|infection prep|of pro:poss:det|your n|computer-PL .

%snd: <02.13.31><02.20.66>

%CYN: so if you open the emails errfr :;:b {you will} * you will just download the virus in your computer errfr ::: so you will get infection of your computers . err_m_m ::: |

*CYN: as for instant communication software like MSN messengers , you may discover in recent years sometimes the MSN messenger will just give you a unknown messengers and got a with a website link .

%mor: prep|as prep|for adj|instant n|communication n|software v|like n:prop|MSN n|messenger-PL pro|you aux|may dis#n|cover prep|in re#n|cent n|year-PL adv|sometimes det|the n:prop|MSN n|messenger aux|will adj|just v|give pro|you det|a n|unknown n|messenger-PL conj:coo|and v|get&PAST det|a prep|with det|a n|website n|link .

%snd: <02.21.24><02.36.09>

%CYN: as for instant communication software like MSN messengers , you may discover errfr ::: in recent years sometimes the MSN messenger will just give you {a} * {a} * a unknown messengers err_s_l :;:a and got {a} ~ with a website link . err_m_s :;:a |

*CYN: if you click the link , you will download the virus file .

%mor: conj:subor|if pro|you v|click det|the n|link pro|you aux|will v|download det|the n|virus n|file .

%snd: <02.36.40><02.40.54>

%CYN: if you click the link errfr :;:b er {you will} * you will download the virus file . errfr ::: |

*CYN: then you will get infection in you computer .

%mor: adv:tem|then pro|you aux|will v|get n|infection prep|in pro|you n|computer .

%snd: <02.40.81><02.43.76>

%CYN: then you will get infection in you computer . errfr ::: |

*CYN: and the last one is to speak is downloading file like MP3 or MVs or movies .

%mor: conj:coo|and det|the adj|last pro:indef|one v|be&3S inf|to v|speak v|be&3S part|download-PROG v|file prep|like n:prop|MP3 conj:coo|or n:prop|MVs conj:coo|or n|movie-PL .

%snd: <02.44.07><02.50.96>

%CYN: and the last one {is} * is to speak is err_s_s ::: downloading file like MP3 or MVs or movies . err_m_m :;;a |

*CYN: though most of the movies may be normal in the website , but it still have probability is that that file should be a virus .

%mor: adv|though qn|most prep|of det|the n|movie-PL aux|may v|be adv|normal prep|in det|the n|website conj:coo|but pro|it adv|still v|have n|probability v|be&3S pro:dem|that det|that n|file aux|should v|be det|a n|virus .

%snd: <02.51.56><03.01.61>

%CYN: mmm er er though most of the movies may be normal in the website errfr :;:b but er {if the} # (0.45) er it still have probability is err_m_s ::: that that file should be a virus . errfr :;:a |

*CYN: if the file is a virus , and you download it , then you get infection in your computers .

%mor: conj:subor|if det|the n|file aux|be&3S det|a n|virus conj:coo|and pro|you v|download pro|it adv:tem|then pro|you v|get n|infection adv:loc|in pro:poss:det|your n|computer-PL .

%snd: <03.01.70><03.06.16>

%CYN: if the file is a virus errfr :;:b and you download it errfr :;:b then you get infection in your computers . errfr ::: |

*CYN: and as for the consequence of the got infection of the virus , to speak in short , it just make you computer can't work properly .

%mor: conj:coo|and prep|as prep|for det|the n|consequence prep|of det|the part|get&PERF n|infection prep|of det|the n|virus inf|to v|speak prep|in adv|short pro|it adv:int|just v|make pro|you n|computer aux|can~neg|not v|work adv:adj|proper-LY .

%snd: <03.06.83><03.17.64>

%CYN: and as for the consequence of the got infection of the virus err_m_s :;;b

(0.59) er (1.04) er to speak in short errfr :;;b it just er make you computer err_m_s ::: can't work properly . err_m_s :;;a |

*CYN: just such as the virus may use up all of your hard disk space .

%mor: adv:int|just qn|such prep|as det|the n|virus aux|may v|use adv|up qn|all prep|of pro:poss:det|your adj|hard n|disk n|space .

%snd: <03.18.37><03.23.94>

%CYN: just such as er the virus {may} * may (0.92) use up all of your hard disk space . err_m_l ::: |

*CYN: make you can't save anything .

%mor: v|make pro|you aux|can~neg|not v|save pro:indef|anything .

%snd: <03.24.42><03.26.47>

%CYN: make you errfr ::: can't save anything . err_m_s :;;a |

*CYN: and make the computer can shut down automatically .

%mor: conj:coo|and v|make det|the n|computer aux|can ?|shut v|down adv:adj|automatic-LY .

%snd: <03.26.57><03.31.49>

%CYN: and make the computer errfr ::: can er shut down automatically . err_m_s :;;a |

*CYN: and use up your RAM and make you can't do anything .

%mor: conj:coo|and v|use prep|up pro:poss:det|your n:prop|RAM conj:coo|and v|make pro|you aux|can~neg|not v|do pro:indef|anything .

%snd: <03.31.91><03.35.65>

%CYN: and use up your RAM errfr ::: and make you errfr ::: can't do anything . err_m_s :;;a |

*CYN: and or just make you can't use some softwares .

%mor: conj:coo|and conj:coo|or adv:int|just v|make pro|you aux|can~neg|not v|use qn|some n|softwares .

%snd: <03.36.18><03.40.54>

%CYN: {and} rpl or just make you errfr ::: can't use some softwares . err_m_s err_m_m :;;a |

*CYN: or make you can't access some others' websites .

%mor: conj:coo|or v|make pro|you aux|can~neg|not n|access qn|some pro:indef|other-PL-POSS n|websites .

%snd: <03.40.68><03.44.36>

%CYN: or er make you errfr ::: can't access some others' websites . err_m_s :;;a |

*CYN: there are many consequence , but I just said those are the main problems in our daily life .

%mor: pro:exist|there v|be&PRES qn|many n|consequence conj:coo|but pro|I adv:int|just v|say&PAST det|those v|be&PRES det|the adj|main n|problem-PL prep|in pro:poss:det|our adj|daily n|life .

%snd: <03.44.75><03.51.30>

%CYN: there are many consequence err_m m ::: but er (0.43) I just said errfr ::: those are the main problems er in our daily life . errfr :;:a |

*CYN: as for how can we deal with the virus infection in computers, I think protection is the the first bet .

%mor: prep|as prep|for adv:wh|how aux|can pro|we n|deal prep|with det|the n|virus n|infection prep|in n|computer-PL pro|I v|think n|protection v|be&3S det|the det|the adj|first n|bet .

%snd: <03.52.05><04.02.03>

%CYN: as for mmm how can we deal with the virus infection in computers err_m_s :;:b I think errfr :::; er (1.02) mmm protection is {the} * the first bet . errfr ::: |

*CYN: it's much better than when when you got virus and to deal with it .

%mor: pro|it~v|be&3S qn|much adv|better prep|than conj:subor|when conj:subor|when pro|you v|get&PAST n|virus conj:coo|and prep|to n|deal prep|with pro|it .

%snd: <04.02.27><04.07.61>

%CYN: er {it's} * it's much better than err_m_s ::: er {when} * when you got {in@} * {in@} * er virus err_m_s :;:a and to deal with it . err_m_s :;;a |

*CYN: so proper protection is easy .

%mor: adv:int|so adj|proper n|protection v|be&3S adj|easy .

%snd: <04.07.93><04.10.48>

%CYN: so proper protection is easy . errfr ::: |

*CYN: you have to choose one good antivirus software and a firewalls of course .

%mor: pro|you v|have inf|to v|choose det:num|one adj|good n|virus n|software conj:coo|and det|a n|firewalls prep|of n|course .

%snd: <04.10.69><04.19.00>

%CYN: {you} * {you can} rpl you have to choose one er good antivirus software and a firewalls of course . err_m_s ::: |

*CYN: and you have to open it when you access the Internet .

%mor: conj:coo|and pro|you aux|have inf|to v|open pro|it conj:subor|when pro|you n|access det|the n:prop|Internet .

%snd: <04.19.45><04.22.41>

%CYN: and {you} * you have to open it err_m_l ::: when you access the Internet . errfr :;:a |

*CYN: and this will reduce the risk simply you got the virus .

%mor: conj:coo|and pro:dem|this aux|will v|reduce det|the n|risk adv:adj|simple-LY pro|you v|get&PAST det|the n|virus .

%snd: <04.23.06><04.29.81>

%CYN: er and (1.35) mmm {this} * this will reduce the risk errfr ::: simply you got the virus . err_m_s :;:a |

*CYN: and other is that you don't to accept the website you don't know .

%mor: conj:coo|and qn|other aux|be&3S pro:dem|that pro|you aux|do~neg|not inf|to v|accept det|the n|website pro|you aux|do~neg|not v|know .

%snd: <04.30.96><04.37.24>

%CYN: and$ er {other} * other is err_m_s ::: that you {don't} * don't to accept the website err_m_s :;:a you don't know . errfr :;:a |

*CYN: and don't open the email you don't know .

%mor: conj:coo|and aux|do~neg|not adj|open det|the n|email pro|you aux|do~neg|not v|know .

%snd: <04.37.77><04.40.75>

%CYN: and don't open the email errfr ::: {you don't} * you don't know . errfr :;:a |

*CYN: and don't download anything ok .

%mor: conj:coo|and aux|do~neg|not v|download pro:indef|anything coo|ok .

%snd: <04.40.92><04.46.02>

%CYN: and {don't} * don't er download anything ok . errfr ::: |

*CYN: and use the instant communication softwares carefully .

%mor: conj:coo|and v|use det|the n|instant n|communication ?|softwares adv:adj|care-FULL-LY^adv:adj|careful-LY .

%snd: <04.46.46><04.51.23>

%CYN: and use the instant er communication softwares er carefully . err_m_m ::: |

*CYN: and just talk to the persons you know .

%mor: conj:coo|and adv:int|just n|talk prep|to det|the n|person-PL pro|you v|know .

%snd: <04.51.72><04.54.23>

%CYN: and just talk to the persons errfr ::: you know . errfr :;:a |

*CYN: don't click unknown things .

%mor: aux|do~neg|not v|click n|unknown n|thing-PL .

%snd: <04.54.53><04.56.27>

%CYN: {don't} * don't click unknown things . errfr ::: |

*CYN: but what an luckily you got virus infection , how can you do .

%mor: conj:coo|but pro:wh|what det|a adv:adj|lucky-LY pro|you v|get&PAST n|virus n|infection adv:wh|how aux|can pro|you aux|do .

%snd: <04.56.92><05.05.19>

%CYN: {and so} # but er (0.52) what an luckily you got {in@} * {in@} * {in@} ~ er virus infection err_s_s :;:b how can you do . err_m_s ::: |

*CYN: I think the first thing is you just open the antivirus softwares to check it to try find the virus in which software or something .

%mor: pro|I v|think det|the adj|first n|thing v|be&3S pro|you adv:int|just v|open det|the n|virus ?|softwares inf|to v|check pro|it inf|to v|try v|find det|the n|virus prep|in rel|which n|software conj:coo|or pro:indef|something .

%snd: <05.05.40><05.15.06>

%CYN: I think errfr ::; the first thing is errfr ::: you just open the antivirus softwares err_m_m :;:a to check it errfr :;;a to try find the virus in er which software or something . err_m_s :;;a |

*CYN: and let the antivirus software to deal with it , to delete it or to leave it alone ok .

%mor: conj:coo|and v|let det|the adj|virus n|software prep|to n|deal prep|with pro|it inf|to v|delete pro|it conj:coo|or inf|to v|leave pro|it adv|alone coo|ok .

%snd: <05.15.59><05.23.20>

%CYN: and let {the} * er the antivirus software errfr ::: to deal with it err_m_s :;;a to delete it errfr :;;a or {to} * (0.50) to leave it alone ok_pfp . err_s_l :;;a |

*CYN: but if this method can't help you , I think the last method is to delete anything including the Windows .

%mor: conj:coo|but conj:subor|if det|this n|method aux|can~neg|not v|help pro|you pro|I v|think det|the adj|last n|method v|be&3S inf|to v|delete pro:indef|anything part|include-PROG det|the n:prop|Window-PL .

%snd: <05.23.73><05.35.76>

%CYN: {and} rpl but er {if} * if {this} * this method can't help you er errfr :;:b I think errfr ::; the last method is errfr ::: to delete anything including the Windows . err_m_l :;;a |

*CYN: or erase all the things and reinstall the Windows .

%mor: conj:coo|or v|erase qn|all det|the n|thing-PL conj:coo|and re#v|install det|the n:prop|Windows .

%snd: <05.36.37><05.40.94>

%CYN: or erase all the things errfr ::: er and reinstall the Windows . errfr ::: |

*CYN: and that will be the final choice because you will lose all of your datas and files .

%mor: conj:coo|and pro:dem|that aux|will v|be det|the adj|final n|choice conj:subor|because pro|you aux|will v|lose qn|all prep|of pro:poss:det|your n|data-PL conj:coo|and n|file-PL .

%snd: <05.41.40><05.48.44>

%CYN: and that will be {the} * (0.63) the final choice errfr ::: ok_pfp because you will lose all of your datas and files . errr_m_m :;:a |

*CYN: so then to remind you , you have to make a back-up frequently even your computer is run properly .

%mor: co|so adv:tem|then prep|to v|remind pro|you pro|you v|have inf|to v|make det|a n|back-up adv:adj|frequent-LY adj|even pro:poss:det|your n|computer v|be&3S n|run adv:adj|proper-LY .

%snd: <05.48.89><05.56.13>

%CYN: so then to remind you errfr :;;b you have to {back up} ~ make a back-up frequently errfr ::: (0.43) er even you computer is run properly . err_m_s :;:a |

*CYN: always back up the important files , photos , anything on some hard base like like DVD-ROMs or another ROM such such as USB stick .

%mor: adv|always adv:loc|back prep|up det|the adj|important n|file-PL n|photo-PL pro:indef|anything adv|on qn|some adv|hard n|base prep|like prep|like n:prop|DVD-ROMs conj:coo|or det|another n:prop|ROM qn|such qn|such prep|as n:prop|USB n|stick .

%snd: <05.56.69><06.13.62>

%CYN: er always back up the important files , photos , anything {on some} * er

(0.94) on some er {hard} * hard base {like} * like DVD-Roms or another er Rom {such} * such as USB stick . err_m_s ::: |

*CYN: this ok .

%mor: det|this adj|ok .

%snd: <06.13.84><06.15.52>

%CYN: er this ok . err_m_s ::: |

*CYN: so save your file carefully .

%mor: conj:subor|so v|save pro:poss:det|your n|file adv:adj|care-FULL-LY^adv:adj|careful-LY .

%snd: <06.15.96><06.20.55>

%CYN: er so mmm (1.44) save your file carefully . err_m_m ::: |

*CYN: and that's all my presentation .

%mor: conj:coo|and pro:dem|that~aux|be&3S qn|all pro:poss:det|my n|presentation .

%snd: <06.21.07><06.22.87>

%CYN: and that's all my presentation . errfr ::: |

@End