TASK AND LANGUAGE PRODUCTION: A STUDY OF TASK CHARACTERISTICS AND TASK CONDITIONS ON SPOKEN LANGUAGE PRODUCTION IN AN ENGLISH LISTENING AND SPEAKING COURSE IN A UNIVERSITY OF SCIENCE AND TECHNOLOGY IN CHINA

XI CHEN, BA, MA

Thesis submitted to the University of Nottingham for the degree of

Doctor of Education

JULY 2022

DECLARATION

I confirm that this dissertation is my original work. It does not include material previously presented for the award of a degree in this, or any other university.

Signed ……

Xi Chen

July 2022

ABSTRACT

This study investigates task characteristics and task conditions in an English Listening and Speaking course to view the varied features in task characteristics and task conditions and their influence on students' spoken language production. The context of the study is set in a College English (CE) course, English Listening and Speaking, in a university of science and technology in China. Following the context, the literature review expatiates on 3 themes: tasks; spoken language production; task and spoken language production. By exploring the key themes, the research question is set as follows.

How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?

For this study, the research methodology is established in the philosophical paradigm, pragmatism, and mixed method research. This study has designed three phrases of research on speech-making tasks with varied task characteristics and task conditions. Both quantitative and qualitative data have been collected and analyzed for varied task characteristics and task conditions and their influence on students' spoken language production.

As for the results, this thesis finds from the quantitative analysis that the familiarity of information and structure can be partly beneficial to the lexical complexity in spoken language production. The task structure can promote syntactic complexity. While the task condition of rehearsal can increase fluency of speech rate. Regarding the qualitative analysis, both questionnaires and focus group interviews results demonstrate that familiarity of topic and task structure, as well as strategic planning

and rehearsal can promote spoken language production. It is accepted by more than half of the online questionnaires' respondents and the focus group participants that with familiar topic, structure, strategic planning, and rehearsal, students could produce better speech in CAF.

Regarding contributions of knowledge, this study provides a reference and supplement for the theoretical framework of tasks and spoken language production in teaching English speaking in a university of science and technology as well as offering teaching practice in the context of this target students for classroom practitioners. Meanwhile, the pedagogical implications in the context of this study can be applied to many different contexts in Chinese ELT.

In terms of implications, first, for CE teachers, this study establishes an opportunity for teachers of CE speaking class to reflect on their teaching practice; to think about how the context of teaching to design and implement more tailored tasks for target students; and create awareness that task design and implementation are influenced by various factors such as familiarity of information, degree of structure, rehearsal, and strategic planning;

Second, the implication for the university of science and technology shows that the General English courses in the CE curriculum should be taught at different levels.

Last but not the least, this study may have implications for government policy makers in China and language teachers in international contexts that would accommodate the learning needs of the diverse students in both China and abroad.

ACKNOWLEDGEMENTS

This Ed. D journey is strenuous. However, I could not reach the end of this journey without the encouragement and support of the following people.

First, I would like to give my heartfelt gratitude to Professor Bob Adamson for supervising me in the research, and in writing this thesis. Without his careful guidance and constant support, the completion of this thesis would not be possible.

Second, I sincerely thank all the faculty members in School of Education and English, Faculty of Humanities and Social Sciences, UNNC, China. Special thanks should be given to Professor Anwei Feng, Dr. John Lowe and Dr. Yanhui Zhang, who kindly provided tutorship of this Ed. D study.

Third, I am grateful for the cooperation of the students in English Listening and Speaking classes in 2020 and the participants of focus group interviews who made this research possible.

Last but not the least, I would like to express my gratitude for the support, patience, and love of my mother, who has helped taking care of my child for 6 years so that I can work on my Ed. D study and the thesis. In the meantime, special thanks should be given to Dr. Joyce Webb and Dr. Xin Gao, who has provided guidance in revising and proofreading the thesis in the final stage.

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
AS-unit	Analysis of Speech Unit
AWL	Academic Word List
CAF	Complexity, Accuracy, Fluency
CAL	Center for Applied Linguistics
CE	College English
CET-4	College English Test-Band 4
CET-6	College English Test-Band 6
CETs	College English Tests
CLT	Communicative Language Teaching
C-unit	Communication Unit
CWR	Content Word Ratio
ELT	English Language Teaching
ESP	English for Specific Purposes
EST	English for Science and Technology
GE	General English
GTM	Grounded Theory Method
L1	First Language
L2	Second Language
MMR	Mixed Methods Research
P1	Phase 1
P2	Phase 2
P3	Phase 3
QUAL	Qualitative
QUAN	Quantitative
SLA	Second Language Acquisition
SPSS	Statistical Product and Service Solutions
TBLT	Tasked-based Language Teaching
TTR	Type-token Ratio

T-unit	Minimal Terminable Unit
UNNC	University of Nottingham, Ningbo, China
WTO	World Trade Organization

LIST OF APPENDICES

Chapter 1 Introduction

1.1 PERSONAL MOTIVATION

This study focuses on the topic of task and spoken language production: specifically, the three major themes of task, spoken language production, tasks and spoken language production. The context lies in teaching English as a second language in a CE course in China. The course is English Listening and Speaking, for first year students, in a university of science and technology. The reasons why I have conducted this research can be explained as follows.

The original purpose of conducting this research can be traced back to 2003. When I was a high school student in my hometown, Zhanjiang, in Guangdong province of China, I was chosen to participate in an English speech contest on the topic of “I Have a Dream”. All the contestants were required to make a three-minute speech about their dreams for the future, in response to the speech made by Dr. Martin Luther King. In my case, I wrote that “I have a dream to become an English teacher”. I was fascinated by English and had been influenced by my family traditions, as both my parents taught Esperanto, an artificial language for international communication. From then on, I embarked on the journey of becoming an English teacher.

In 2006, I began my undergraduate study as an English major at Guangdong University of Foreign Studies in Guangzhou, in southern China. There, I was engaged in the study of British and American Literature. My favorite courses were: Selected Readings in British and American Drama, Poetry, and Short Stories. In the short stories course, I was most impressed by “The Fall of the House of Usher” by the American writer Edgar Allan Poe. Our teacher described Usher’s face and pointed out that the “luminous eyes” from the protagonist Usher could indicate that he was going

through some mental disturbance, which fits the character in the story. For me, I felt so inspired at that time because one simple description of face could show the psychological state of the protagonist.

In the meantime, I gained teaching experience as an English teacher in a language training center for young learners. I taught basic English vocabulary by having the students read and spell the words aloud. However, the drills in reading and repetition of definitions did not prove to be effective in the class for young learners. The students were neither interested or motivated. I was bewildered regarding how to teach an English class which could capture the students' attention and enhance motivation of English learning.

In order to seek further education in English teaching, I applied for postgraduate study at the Center for Applied Linguistics (CAL) at The University of Warwick in England, hoping to equip myself with the pedagogical knowledge of an English teacher. In Warwick, I majored in English Language Teaching (ELT) and had a growing interest in the ELT approach, Task-based Language Teaching (TBLT). TBLT is student-centered and it can enhance students' motivation and participation in class (Fang, 2003; Zeng, 2007; Zheng, 2010). One major module in CAL was called the ELT Methodology. In class, teachers introduced different English teaching approaches and asked us to choose one to develop an English lesson in 20 minutes and teach the class to our peers. I was inspired by the combination of TBLT and literature which uses literature as the contents of learning and TBLT as the approach to teaching. Therefore, I designed a task-based class on Maupassant's *The Necklace* to teach vocabulary and speaking. After the first peer teaching practice, my peers reflected that the class was too form-focused and there was not enough time for speaking practice. Then, I developed the second peer teaching based on the story "The Fall of the House of Usher" by Edgar Allan Poe. The second attempt received more positive feedback from my peers. One reviewer told me "This lesson is a strong form of TBLT of

multiple meaning-focused tasks to help students immerse in the story and also practice speaking". From that time, I became fond of task design and was eager to know how tasks could be implemented in class. It ignited the passion that eventually resulted in this study.

Later in 2012, when I came back to China to work as an English teacher at a university of science and technology, the first course I taught was Academic English Writing for postgraduate students. The course contents were mainly about writing process, essays, and research proposal. For this course, it was not easy to be interesting and stimulating. Then, I attempted to apply task-based approach to encourage meaningful communication and group cooperation. For example, when talking about the process of academic writing, I created a task for students to design a poster to express their opinions on the actual process of academic writing. For the task performance, most of them were motivated by the task.

In 2014, I was able to construct a general elective course. I immediately thought of British and American Literature as I reveled in reading literature and I wanted to apply TBLT in this course to practice task design and implementation. However, when I designed a speaking task in the literature class, some of the students thought that it would be very demanding, because speaking was a weak link for most of them as second language (L2) learners. In the meantime, they need to comprehend the literary works and articulate the information. For instance, when I asked them to perform role playing of Romeo and Juliet to practice their speaking, the students considered it very challenging. Another problem for the teachers also came with the implementation of the role play task to enhance their spoken English.

Consequently, these problems triggered reflection as to what changes could be made in the classroom regarding how to design and implement role playing tasks to enhance students' spoken English. By reviewing previous research, the design of a task in

literature with respect to the task characteristics remain unknown (Gu & Jin, 2021). The impacts of some task characteristics (familiarity of information; degree of structure) on students' spoken language production were unclear to classroom practitioners (Skehan, 2014). Even though, it is specified in the General Guidelines of College English Teaching (Ministry of Education, 2020) that the task-based approach can be applied in the CE classroom, teachers still had little idea of how to implement tasks with the task conditions (strategic planning; rehearsal) to facilitate students' spoken language production (Ellis, 2013). There were few research findings or pedagogical experiences available, especially in universities of science and technology in China (Gu & Jin, 2021).

Therefore, I was attracted to filling in the above research gaps when I undertook this doctoral program. As a result, this thesis investigated the following research question: *How do different task characteristics and task conditions impact students' spoken language production (CAF) in an English Literature course?*

The original task type devised for this study was a role-playing task in the elective course of British and American Literature class, in the Spring of 2020. The data was designed to be collected at that time. However, this intended arrangement of research was interrupted by COVID-19 in 2020. Because of the spread of the pandemic, the spring semester of 2020 was changed to online teaching, which had made the role playing in groups impossible as students were in different cities in their hometowns to take online courses. Consequently, data could not be collected at that time.

Taking the successful completion of this thesis into consideration, I needed to switch to a flexible plan. From 2016 on, I had taught the General English course: English Listening and Speaking, for undergraduates in my school every year. TBLT was also applied to this course. To assess students' spoken English, students were required to make a speech on a given topic in the language lab at the end of each term as the

speaking test for final exam. But the final speech performance was not satisfying. Most students could not produce accurate, fluent, and complex language. For a few years, I was perplexed about how to design the speech task and implement it in class. How can teachers design speech tasks with varied task characteristics and implement them with different task conditions so that students could have better spoken language production?

Bearing these in mind, I discussed with my supervisor, Professor Bob Adamson, to change task types, courses, and target students. If I had stayed with the earlier design, I would have to wait for one more year for the elective course to be reopened in 2021, which could postpone the thesis completion process. Thus, the design of this study was changed to speech-making as the task type, first year students as the target, and English Listening and Speaking as the course. This class was available in the fall of 2020 and could guarantee timely data collection. After explaining the situations to my supervisor, approval was granted for solving this problem.

Going through the twists and turns, the final research question of this study is as follows.

How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?

The task characteristics of familiarity of information, degree of structure in the task design phases as well as the task conditions of rehearsal and strategic planning in task implementation have been investigated in English listening and speaking to view their influence on students' spoken language production. With the findings in minds, teachers in this course can design and implement the speech tasks to promote students' spoken language production.

1.2 CONTEXT OF THE STUDY

English Learning in China

Ever since the open-door policy of the 1980's, English learning in China gained a significant boost in international trade and tourism along with the increase of well-paid jobs (Adamson, 2004a). Entering the millennium, English in China has gained its crucial importance in globalization including important internationalized events: the Shanghai APEC meeting in 2001, the Beijing Olympic Games in 2008, the Shanghai World Expo in 2010. Along with its influence in international exchange and communication, English in China today has reached its high point in education. In formal education, students can learn English in kindergarten in some cities from songs, games, and toys. The official announcement from The Ministry of Education in 2001 indicated that ELT begins in grade three, in primary school, supporting a view to start in grade one. ELT continues in the junior and senior secondary school (Gil & Adamson, 2011).

Moving to the university level, English language courses are required for every student. ELT in universities diverges in two directions: one for English majors and the other for non-English majors (Cortazzi & Jin, 1996; Yao, 1993). For this study, the English courses for non-English majors, College English, are focused on (Wang, 1999; Gil & Adamson, 2011). Since the focus of this study is General English course, English Listening and Speaking, in CE, the CE curriculum is introduced.

Regarding the CE curriculum in higher education in China, three branches are set out from the General Guidelines of College English Teaching (Ministry of Education, 2020). The three levels of CE teaching are General English, English for Specific Purposes (ESP) and Intercultural Communication courses (Wang & Wang, 2019). The General English courses are designed as a fundamental part of CE curriculum with the purpose of developing language skills in English listening, speaking, reading, writing and translation to increase students' basic knowledge of society, culture, and science to broaden international vision (Ministry of Education, 2020). The ESP courses focus

on the field of English use, with the goal of enhancing students' ability to use English for professional and academic communication and work (Ministry of Education, 2020). The Intercultural Communication courses aim to help students understand the differences between Chinese and foreign world views with the goal of enhancing intercultural awareness and intercultural communication competence (Ministry of Education, 2020). For this study, the context lies in the General English course of English Listening and Speaking.

Concerning teaching methods, the CE teaching method is suggestive to be more comprehensive and innovative, and it is agreed that CE teachers can tailor their teaching and adapt various teaching methods in their contexts (Li, Xing & Wang, 2019). According to the General Guidelines of College English Teaching (Ministry of Education, 2020: 33), TBLT is advocated in CE on the national level, as it is suggestive that the CE teaching can adopt the task-based method.

> College English teaching can adopt task-based, cooperative, project-based, inquiry and other teaching methods. The teaching activities should change from "teaching" to "learning", to form a teaching process guided and inspired by teachers and characterized by students' active participation.

However, problems have arisen in practical perspectives. For CE teachers, the design of task-based teaching classroom has been more challenging than the traditional one (Li, 2020). Teachers need to consider the difficulty of tasks and the impact of various variables (input, learning, and procedural factors) on the tasks. Therefore, when teachers design a task-based syllabus, it can become very demanding. The implementation of a task is hard to predict for the actual outcomes. Besides, it is problematic to execute TBLT in class (Li, 2020). During the implementation of the task, some students will use their mother tongue instead of the target language due to the limited vocabulary and grammar knowledge. Meanwhile, the large number of students in an English class in China makes it difficult to implement task-based teaching method (Li, 2020).

Accordingly, based on the above personal motivation and context, this study has striven to address the problems in the task design and implementation level on students' spoken language production.

1.3 THEMES OF THE STUDY

This research focuses on three major themes: tasks, spoken language production, task and spoken language production: the relationship between task characteristics and spoken language production, task conditions and spoken language production. For the first theme, tasks, the definition, and task types will be explained.

1.3.1 Definition of Tasks

Authors	Task Definitions
Prabhu (1987: 24)	(A task) is an activity which requires learners to arrive at an **outcome** from given information through some process of thought, and which allows teachers to control and regulate that **process**.
Candlin (1987)	(A task is) a set of differentiated, sequencable, problem-posing activities involving learners and teachers in some joint selections from a range of varied **cognitive and communicative procedures** applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent **goals** within a social milieu.
Willis (1996: 23-25)	A task is always the activity where the target language is used by learners for a **communicative purpose** to achieve an **outcome**.
Skehan (1999)	Meaning is primary; The assessment of the task is in terms of **outcome**.
Samuda & Bygate (2010: 16)	1) It involves holistic language use; 2) It requires a meaningful target **outcome;** 3) It necessarily involves some individual and group processes; 4) It depends on there being some input material.
Ellis (2013: 2)	6 dimensions should be considered: **scope, perspective, authenticity, linguistic skills, psychological processes, and outcomes**. Tasks should be viewed based on their scopes such as workplans, perspectives focusing on meaning, authenticity of real-world processes, or any of the 4 linguistic skills, the cognitive processes, and the communicative outcomes.

Table 1.1 A List of Task Definitions by Researchers and Scholars

After reviewing various definitions of researchers and scholars, tasks are defined in Table 1.1. Based on the keywords from the below list, conclusions can be drawn on the definition of a task. A task is goal-oriented (Candlin, 1987), with a number of steps, which follows a series of cognitive and communicative procedures, and has defined outcomes (Prabu, 1987; Candlin, 1987; Willis, 1996; Skehan, 1999; Samuda & Bygate, 2010; Ellis, 2013). Additionally, a task is sequential and can be subject to pedagogical interventions (Prabu, 1987; Candlin, 1987).

1.3.2 Task Types

The task types, which focus on form and focus on meaning have been distinguished. Long (1991: 44) has explained that "focus on form" concentrates on "the contents of syllabus and lessons based on the linguistic items. "Focus on meaning" requires learners to concentrate on the meaning they want to express (Zhang, 2006).

By taking a more comprehensive view on the use of tasks, Littlewood (1992) regarded the tasks with the focus on form and meaning, as a continuum with focus on language as a medium, on one end, and focus on learners' message on the other. In addition to the weak, medium, and strong categories of tasks from Littlewood (1992) and Legutke and Thomas (1991), a new framework of the types of tasks in English language teaching was adapted (Tong et al., 2000) (see Figure 3.2).

In this research, the target students were first year college students. Most of them had received sufficient training in English for the College Entrance Exam. However, their speaking performance was not satisfactory. Therefore, the strong tasks with purposeful, authentic communication, and open-ended outcomes were chosen, as this study sought meaningful and real-life communication from the learners.

Speech-making Tasks

In this study, the speech-making tasks are chosen to ensure purposeful, authentic communication, and open-ended outcomes. In the classroom practice, meaningful and real-life communication from the learners is practiced through brainstorming for ideas and negotiating meanings.

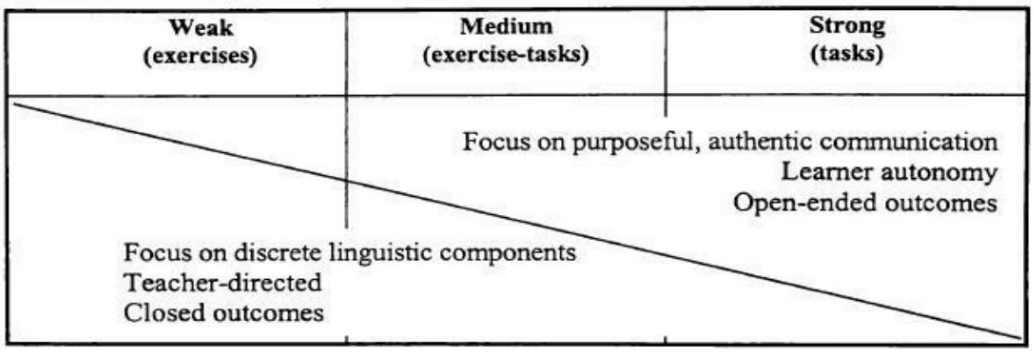


Figure 3.2: Types of Tasks (adapted from Tong et al., 2000)

1.3.3 Spoken Language Production

For first language production, Levelt (1989) proposed the information processing model in three hierarchically modular stages: conceptualization, formulation, and articulation. The conceptualization stage is to develop and organize ideas to a communicative goal. Then, in the formulation stage, a phonetic plan is made for the content of speaking. In the end, articulation is created when the phonetic plan is transformed into actual speech (Ellis, 2013).

This study also gives focus to second language production which is not as simple as the first. To fully understand second language production, learners' second language acquisition in the Input and Output Hypotheses has been clarified. Based on the Output Hypothesis, the triad constructs of complexity, accuracy, and fluency were identified by Skehan (1999), which can be served as the measurements of second language speech production.

In second language acquisition, Krashen (1982) emphasized that language input is the most important in the Input Hypothesis. What learners need is the comprehensible input, which means that the language input should be understandable. The Second Language Acquisition (SLA) will occur when learners can understand the input a little beyond their language level (Krashen, 1982). Therefore, learners can acquire language and understand the message in the target language when they are exposed to language that is slightly beyond their actual level. If the current level is "i", then the acquirer can understand input that contains "*i + 1"* (Krashen, 1982; Nunan, 2011).

However, several studies have shown that the "comprehensible input" alone is not enough to make second language learners achieve high levels in SLA (Swain & Lapkin, 1995). To supplement the inadequacy of the Input Hypothesis, Swain (1985) put forward the Output Hypothesis, which suggested that language engaging in the syntactic process of the learners could foster SLA (Ellis, 2013).

1.3.4 CAF Constructs

Based on Swain's Output Hypothesis, three perspectives of second language production have been distinguished by Skehan (1999). Complexity, Accuracy, and Fluency (CAF) have been viewed as the principal research variables of language production in L2 research (Skehan, 1999; Ellis, 2009). The three constructs of second language production are defined in Table 1.2.

Complexity	The capacity to use more advanced language, with the possibility that such language may not be controlled effectively. This may also involve a greater willingness to take risks.
Accuracy	The ability to avoid errors in performance, possibly reflecting higher levels of control in language, as well as a conservative orientation.
Fluency	The capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems.

Table 1.2: Definitions of Complexity, Accuracy, and Fluency (Adapted from Skehan & Foster, 1999: 96–97)

The L2 pedagogy research on the three constructs could be traced back to the 1980s. Brumfit (1984) was one of the earliest researchers to identify the dichotomy between fluency-oriented and accuracy-oriented activities. Later in the 1990s, Skehan (1989) introduced a third component of the triad, complexity, to form the CAF (complexity, accuracy, and fluency) in the proficiency dimensions. Accuracy relates to the "degrees of deviations from a particular norm" (Housen & Kuiken, 2009: 3). Accuracy, compared with errors, is characterized as a deviation from form (Housen & Kuiken, 2009). Fluency is about the language proficiency of a person with characterization of ease, eloquence, and smoothness of speech (Housen & Kuiken, 2009). The last of the most ambiguous triad is complexity, which can be both the "properties of language task and the properties of L2 performance and proficiency" (Housen & Kuiken, 2009: 3).

1.3.5 Task Characteristics and Spoken Language Production

As for the relationship between task characteristics and spoken language production, Skehan (2014) has identified the tasks features including familiarity of information (concrete-abstract; familiar-unfamiliar materials) and degree of structure (structured-unstructured) in the task design process. For the familiarity of information, the tasks "vary as to whether they require information that is familiar to the participants as part of their personal experience" (Skehan, 2011: 235). For the degree of structure, some tasks have a clear and over-arching structure while some do not (Skehan, 2011). In this study, the relationship of familiarity of information and structure with spoken language production in the research context are scrutinized.

1.3.6 Task Conditions and Spoken Language Production

In task implementation, the task conditions can be divided into three phrases: pre-task planning, during task processing and post-task activities (Skehan, 2014). Pre-task

planning is related to task preparedness, such as planning and repetition. During task processing involves task processing of time pressure and visual support. Post-task activities include the activities and exploitation after the completion of tasks (Skehan, 2014). In this study, the speech-making task has been selected. Regarding this task, the pre-task planning phrase has been the focus, as teachers can apply pedagogical interventions in the task preparation process.

Pre-task Planning

The pre-task planning phrase, which is the planning before learners perform a task, is identified by Ellis (2005) and divided into two aspects: rehearsal and strategic planning (see Table 1.3). In this study, the relationship of rehearsal and strategic planning with spoken language production has been investigated.

Pre-task planning is planning that is done before learners perform a task.	Rehearsal	Planning takes the form of an opportunity to perform the complete task once before performing it a second time.
	Strategic planning	Planning includes the contents to be expressed and the language to be used but without an opportunity to rehearse the complete task.

Table 1.3: Pre-task Planning and Types of Pre-task Planning (Ellis, 2009: 474)

1.4 AN OVERVIEW OF METHODOLOGY

This study has adopted a Mixed Method Research (MMR) of both quantitative and qualitative techniques (Johnson & Onwuegbuzie, 2004). As for the mixed method design, there are three major decisions in determining a mixed design method (Creswell & Plano-Clark, 2007). The first is the timing of quantitative and qualitative designs. This differs from concurrent and sequential timing. The second is weight which determines the importance of quantitative and qualitative method and makes clear whether they are equal or unequal. Finally, the third is the mixing of quantitative and qualitative methods (Creswell, et al., 2003).

For this research design, the timing of quantitative and qualitative methods is

sequential with quantitative method as the focus and qualitative as the supplement. The weight of importance of the quantitative and qualitative designs is unequal and there is a merging of data analysis of the two methods. The explanatory model of triangulation has been selected to conduct this research based on the above three decisions. The CAF measures are specified in the conceptual framework of this study. The quantitative data of CAF measures have been collected for speech tasks. The qualitative data of questionnaires and focus group interviews have been analyzed for students' perceptions on the varying features in task characteristics, task conditions, and their influence on spoken language production. The quantitative data were collected before the qualitative data. Meanwhile, both data were merged to analyze and interpret the research findings to provide comprehensive perspectives in viewing the research issues as well as increasing the validity of the data (Creswell, 2014). The quantitative data can produce numeric findings for the influence of task characteristics and task condition on spoken language production. In the meantime, the qualitative data has probed for deep insights from students' viewpoints of the research question (Creswell, 2014).

1.4.1 Research Design

Three classes of students from English Listening and Speaking were research participants in this study. Each class had about 40 students. Altogether, there were approximately 120 students. One convenience sample group of 6 students was selected for focus group interviews.

This research plan incorporated three phases. Phase 1 began from Week 7 of the course in the fall of 2020. Every student made a speech based on a familiar topic with strategic planning but not the structured outline. Phase 2 was conducted one week later in Week 8. The phase 2 task was unfamiliar but not structured with rehearsal and strategic planning. After phase 1 and 2, the online questionnaires for the 3 classes and

focus group interviews were collected. Phase 3 was one week after phase 2. The phase 3 speech task was familiar, structured with strategic planning and rehearsal. After phase 3, online questionnaires for the 3 classes and the focus group interviews were obtained. All the speeches in the 3 phases were recorded.

1.4.2 Data Analysis

For quantitative data, the recordings of 30 participants were analyzed in SPSS (Statistical Product and Service Solutions) for statistical analysis of the research question. For qualitative data, online questionnaires were analyzed along with focus group interviews. The online questionnaires were conducted with keywords analysis. The focus group interviews were analyzed by the grounded theory approach for insightful answers from the students on task characteristics and task conditions and their influences on spoken language production.

However, problems arose in the quantitative data analysis process. In order to obtain findings in the quantitative data, complicated calculation needed to be done for the numbers of the CAF measures. Furthermore, SPSS statistical consultation and training were taken to determine the answers to the research question. For the consultation sections, online appointments were made with Dr. Yanhui Zhang, an expert in SPSS analysis in UNNC (University of Nottingham, Ningbo, China). From Dr. Zhang's suggestions, the sample size increased from the original 6 students' recordings of 30 MP3 speeches to 30 students with 150 MP3 recordings. The listening and transcription process took approximately three weeks to complete.

1.5 ARGUMENT

This study contributes to the theoretical framework of tasks and spoken language production in teaching English speaking. It argues that familiarity of topic, structure

and rehearsal are partially related to the promotion of spoken language production, which can be regarded as a reference in the future practice and study for related research. In terms of context (English Listening and Speaking under the CE curriculum), the thesis provides valuable pedagogical insights that can be applied to many different contexts in ELT. Additionally, the results of this study can be of value for future practitioners and researchers who would like to apply TBLT in the speaking class either in their teaching practice or future research.

1.6 OVERVIEW OF THE THESIS

This chapter has explained the personal motivation of this study. Further, the context and the research question have been established. Then, an overview of the themes of this research has been created, which is about task and spoken language production as well as the relationship between task characteristics and spoken language production, task conditions and spoken language production. Next, the overview of methodology including the research method, research design, and data analysis has been specified. Finally, the findings of this research have been explained, and the significance of this study has been determined.

Moving to the layout of the thesis, Chapter 2 reveals the context of the study, with focus on the course of English Listening and Speaking under the CE curriculum in a university of science and technology in China. Following the context, Chapter 3 depicts the framework for literature review, which is about tasks, spoken language production, task characteristics, task conditions, and their influence on spoken language production. Chapter 4 lays out the methodology section with the philosophical paradigm, pragmatism, and selected, mixed method research design. Chapter 5 illustrates data analysis, including: SPSS statistical analysis of quantitative data, key words analysis of qualitative questionnaires, and the grounded theory approach for focus group interviews to determine the varied task characteristics and task conditions and their influence on students' spoken language production. Chapter

6 concludes the thesis with discussion and implication for further research. In the next chapter, the exploration of context will be displayed.

Chapter 2 Context

This section introduces the background and context of the study from English as a global language perspective to English language teaching (ELT) in China. Then, it is narrowed down to College English (CE) teaching in China. The context of the study is set within the scope of teaching speaking in CE courses in a university of science and technology. For pedagogical practice, the traditional teaching method is problematic in the speaking course. Therefore, task-based language teaching (TBLT) has been adopted. However, challenges also exist for language teachers. The impact of task characteristics, task conditions in spoken language production remain unknown to teachers. To explore these unknown issues, the research question and the context of this study will be presented in this chapter.

2.1 ENGLISH AS A GLOBAL LANGUAGE

2.1.1 The Rise of English as a Global Language

English has risen as a global language in the past two centuries. Crystal (1997) defined a global language to be one which gained global status and worldwide recognition by every country. To achieve such a global status, one language should have the political, military, and economic powers from its country to establish and maintain its status (Crystal, 1997). English, however, has achieved the special role of recognition of global status by other countries “in the right place at the right time” (Crystal, 1997: 78). Historically speaking, Britain became the world’s leading power in industry and trade in the 19th century, along with colonialism and imperialism, to spread English around the world. After the 20th century, America obtained its economic superpower and supremacy, which maintained the global status of English (Crystal, 1997). Nowadays, countries, all over the world, are connected to one another through exchanges and communication in various fields, such as economy, politics,

technology, and education. As English has already been entitled with linguistic power in all fields, it is now needed as a means of global communication to ensure the effectiveness of these exchanges (Hu, 2007). Therefore, English, as a global language, has maintained its global status and international recognition by other countries.

2.1.2 The Three Facets of Global English and China's Position in Kachru's Three-circle Model

By differentiating the special status and recognition accepted by every country, Crystal (1997:3-4) has distinguished three different facets of "global English". The first comes with the countries such as America, Britain, and Canada with whose people speak English as the first language. The second facet refers to those countries such as Ghana, Nigeria, and India who have gone through colonization and established English to be the official language or second language, which is used as "the medium of communication in the official domains" (Crystal, 1997: 4). The third level includes those countries such as China, Russia and Spain which prioritize English in foreign language teaching even though it is not the official language. Sharing similar views, Kachru (1985) has developed the three-circle model of English (see Figure 2.1). The inner circle refers to those people whose first languages are English. The outer circle includes those non-native English-speaking countries who consider English as an institutionalized and official language. The expanding circle refers to those countries like China and Russia who choose English as a foreign language in education and teaching. Moving our angles towards the expanding circle, English as a global language, is being widely taught and learned in China (McArthur, 2004).

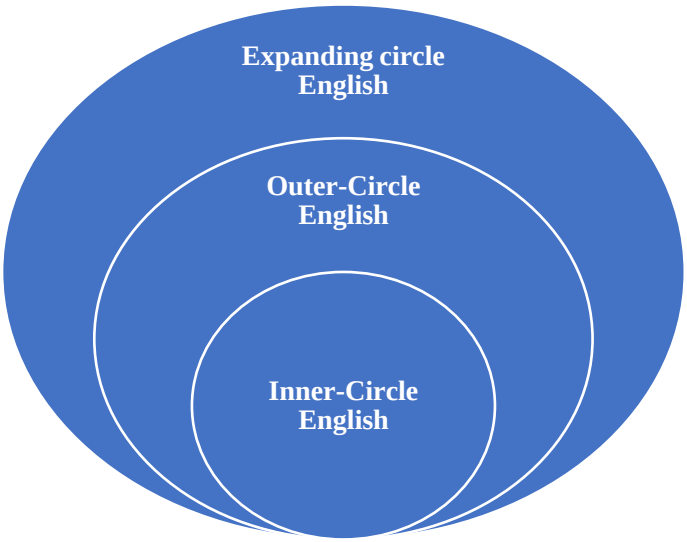


Figure 2.1: Kachru's Three-circle Model of English

2.2 ENGLISH LANGUAGE TEACHING (ELT) IN CHINA

When considering ELT in China, China's early interactions with English in relationship to ELT in the 20th century must first be explored. This study will, then, focus on ELT in millennium China.

2.2.1 China's Early Interaction with English (1759-1911)

Situated in the expanding circle, China' s early interaction with English and ELT can be traced back to the late Qing Dynasty (Adamson, 2004a), during which the status of English had shifted from a barbaric tongue to a powerful language. The earliest interaction between China and English began in the mid-nineteenth century, when trading empires such as Britain and America sought entrance to the market in China. Christian missionaries attempted to infuse Chinese minds (Adamson, 2004a). Large numbers of Chinese had studied English and many became Christians (Adamson, 2004a). However, the Emperors in China, at this stage, considered themselves, in the Celestial Empire while the other countries, as the terrestrial domains. At that time, only the nominated port Guangzhou was open for foreign trade. English was then

regarded as a barbaric tongue. Nevertheless, Chinese pride in the sluiced-door policy was shattered in reality. The feudalistic society was beaten by the advanced technology and modernization of European countries (Adamson, 2004a). The Opium Wars had defeated the Celestial dream in China and the wounded Chinese were forced to learn about Britain and its language. English was later treated and taught as a powerful tool to access western scientific knowledge and to develop international diplomacy. Furthermore, it was endowed with wealth in jobs and powerful status after the Opium Wars (Adamson, 2004a).

2.2.2 English Language Teaching (ELT) in 20th Century China

Moving toward the 20th century, ELT has experienced its ups and downs along with the historical development in China. During the Republican Era (1919-1949), English was taught as a vehicle for the exploration of western philosophy and intellectual communication with the west (Adamson, 2004a; Liu, 2020b; Zhou, 2019). Since the founding of The People's Republic of China (1949-1960), foreign language teaching in China was under the influence of the Soviet Union. In the 1950s, Russian had become the official recommendation of foreign language to study in China. No English textbook was produced at that time. Book importations from English-speaking countries were forbidden (Adamson, 2004a). Therefore, English language teaching and learning were discontinued. Later, however, international situations changed dramatically. The early 1960s witnessed the Sino-Soviet split, which directly caused the renaissance of English language learning and the downfall of Russian. English learning was considered as an access to higher education. Emphasis was changed again to the quality of ELT (Adamson, 2004a; Liu, 2020b). Nonetheless, the booming of English teaching and learning was like a flash in the pan. It was dismantled and forced to cease in the early phrase of the Cultural Revolution (1966-1976). The teaching and learning of English were associated as taking the capitalist road. English language teachers were "criticized, vilified, and persecuted". Teenage

students were relocated to rural areas far away from home for physical labor. ELT was suspended all over China (Adamson, 2004a: 107-109; Liu, 2020b). However, it was resumed with Mao's statement to "encourage the young to learn English early". But the English teaching and learning were "highly politicized" (Adamson, 2004a: 109; Liu, 2020b). It was not until the end of the Cultural Revolution that China had regained revitalization of the country's economy and education development. Under the leadership of Deng Xiaoping, the fundamental task in China, at that time, shifted from the Cultural Revolution to economic development to achieve the four modernizations. Special Economic Zones in coastal regions were later established as "the pioneers of economic modernization". The Reform and Opening-up policy was then initiated and enacted to support international trade and exchange (Adamson, 2004a: 129-130; Liu, 2020b). In the mid-1980s, the economic development in China witnessed the resurgence of ELT and brought along a significant boost and fascination for the English teaching and learning (Guo, 2016).

2.2.3 English Language Teaching (ELT) in Millennium China

Entering the millennium, English in China has gained its crucial importance in globalization. Important internationalized events including the Shanghai APEC meeting in 2001, the Beijing Olympic Games in 2008, and the Shanghai World Expo in 2010 were held. The entire nation has acknowledged the significant values in learning English, which can lead to well-paid jobs and opportunities in "international business, study, and tourism" (Adamson, 2004a: 169). Throughout the country, there was unprecedented passion of English learning. Along with its influence in international exchange and communication, English in China, today, has reached its high point in education. In formal education, students can learn English in kindergarten in some cities from songs, games, and toys. The official announcement from the Ministry of Education in 2001 described that the English language teaching begins in grade three in primary school, supporting a view to begin in grade one. The

English language teaching continues in the junior and senior secondary school (Gil & Adamson, 2011). The “Double Reduction” policy was enacted in 2021. It boosted the transition of basic education (primary, junior, and senior secondary school) to focus on the improvement of education quality. English teaching in basic education should take improving the quality of classroom teaching as the priority (Ministry of Education, 2021). Moving toward higher education, English has long been the prerequisite for college entrance exams all over the country (Adamson, 2004a).

2.3 ENGLISH LANGUAGE TEACHING (ELT) IN HIGHER EDUCATION

Ever since the millennium, English, with its “international stature” (Lam, 2002: 246-7), played a vital role in higher education in China (Gil & Adamson, 2011). According to the Quality Report of Higher Education in China (Ministry of Education, 2016), the number of students in higher education by the year 2015 was 37 million, ranking first in the world. For every student in higher education, English is a required and compulsory course.

2.3.1 English as a Major and Non-major in Higher Education

For ELT in the higher education, the directions are divergent in two aspects: one for English majors and the other for non-English majors (Cortazzi & Jin, 1996; Yao, 1993). For this study, College English for the non-English major will receive focus (Wang, 1999; Gil & Adamson, 2011).

2.4 COLLEGE ENGLISH (CE)

As mentioned above, ELT in China has gained revitalization ever since the Reform and Opening-up in 1978. Next, the development of CE will be examined with analysis of distinguished periods, teaching methods, and teaching orientations (Table 2.1).

Periods	Teaching Methods	Teaching Orientations
The Recovery Period (1978-1995)	Grammar-Translation Method	General English
The Development Period (1996-2001)	Grammar-Translation Method	English for Literacy
The Reform Period (2002-now)	Diverse teaching methods: Task-based language teaching; Cooperative language teaching; Experimental language teaching	Emphasis on Listening and Speaking

Table 2.1: College English Since 1978 (Adapted from Li, 2019; Li, Xing & Wang, 2019)

2.4.1 The Recovery Period (1978-1995)

Teaching Method: Grammar-Translation Method

The dominant teaching method at this stage (1978-1995) of CE teaching was the Grammar-Translation Method, as it was still influenced by the concept of reserved and elite education in the early stages of Cultural Revolution (Li, 2019). The CE teaching still focused on the language knowledge itself and concentrated on the reading skills. Although new teaching concepts and methods, such as Communicative Language Teaching, were brought in from abroad, CE, at this time, was basically teacher-centered and exam-oriented. The learning process was mainly based on mechanical memory (Li, 2019). Therefore, the Grammar-Translation Method occupied an important position in teaching practice (Li, Xing & Wang, 2019).

Teaching Orientation: General English

The debate of the teaching orientation between General English (GE) and English for Science and Technology (EST) were going on at this stage (1978-1995) with General English becoming victorious.

The academic community, at this stage, had begun a debate between General English (GE) and English for Science and Technology (EST). Special attention was paid to EST, but there were few advocates for EST (Cai, 2015). Therefore, arguments on the

teaching orientations were on between teaching GE and EST.

In the National Forum on Foreign Language Education in 1978, Deng Xiaoping had pointed out that “the development of modern economy and technology calls for the rapid improvement of the quality and efficiency of education” to support the realization of four modernizations (Cai, 2015: 36). Later, from a commentary review published in *People's Education*, it was specified that CE education should educate talents of science and technology who understand both their majors and the foreign language (Cai, 2015).

Even though there was the specification of EST, dissenting opinions had existed in the field of CE to emphasize GE instead of EST. The GE supporters held the opinion that ELT in universities should mainly enable students to master the basic and common language skills. While those who supported EST believed that the main purpose of ELT in China was to solve the problem of professional readings, especially in science and technology. The professional English language teaching could also pave the way for English learning (Cai, 2015). In 1980, with the support of the National Education Committee, the first CE Syllabus, since the Cultural Revolution, was officially promulgated, which positioned CE teaching in scientific and technological English (Cai, 2015). However, opponents of EST argued that even though the English level of college students had improved, it was not good enough to pass the examination. The practice of seeking instant success and benefit in scientific readings was temporarily effective and had many shortcomings, in the long run. Meanwhile, learning the English language was inseparable from mechanical memory, reading, and recitation. Therefore, due to the limitation of science textbooks, EST was not conducive to the development of English learning, especially to developing communicative competence (Cai, 2015). In the 1980s, the battle between EST and GE ended in victory for the latter.

The success came with the implementation of the CET-4 (College English Test-Band 4) and CET-6 (College English Test-Band 6) nationwide by the National Education Committee in 1987. As CET-4 and CET-6 were based on ordinary and general English, it was also stipulated that all college students must take part in the CETs (College English Tests) after they finish the CE curriculum. Under the pressure of CET-4 and CET-6, universities all over the country had chosen General English, so that the teaching orientation of General English in CE, at this stage (1978-1995), was finally determined (Cai, 2015).

2.4.2 The Development Period (1996-2001)

Teaching Method: Grammar-Translation Method

In the mid-1990s, Grammar-Translation Method was still the principal teaching method even though there were increasing needs for multiple teaching methods in CE teaching.

At that time, China began to actively bid for the entry to the WTO, being eager to participate and integrate into political, economic, technological, and cultural communication all over the world. Nevertheless, the biggest challenge was the lack of people's English communicative ability (Cai, 2015). Therefore, to meet the growing communicative demands in teaching, opportunities were created in communication and the practical use of English (Li, Xing & Wang, 2019). Various teaching methods were encouraged to promote communication and the use of language. However, CE teaching, at this phase, was heavily influenced by traditional teaching methods and the influence of CET-4 and CET-6. In the actual classroom teaching, "Grammar-Translation Method" was still the mainstream teaching method. The discussion on the teaching of language was mainly limited to the reading and vocabulary (Li, Xing & Wang, 2019).

Teaching Orientation: English for Literacy

To resolve the dilemma, debates of CE teaching were held between English for literacy and English for practical use, with the becoming former prevailing. Cen Jianjun, director of Foreign Language Department in Minister of Education in 1999, claimed that CE teaching did not follow the pace of economic and technological development, although ELT had made some progress. Enterprises and employers required college students not only to have the ability of reading and writing, but also speaking and communication. However, the college students tended to speak "dumb English". They could achieve high scores in English tests and exams, but had low ability in speaking and communication (Cai, 2015). Most scholars and teachers seconded the opinion: "our students learn English from primary schools all the way to the graduate schools. However, most people are still unable to communicate in English. Students need to improve listening and speaking" (Cai, 2015: 38).

Nonetheless, this debate had not changed the orientation of CE teaching. The 1999 CE syllabus still revised the syllabus according to its own plan, putting reading as the priority. Although it was found in the survey of social needs that society had a strong demand for listening and speaking, "reading as the first level of requirement" in the 1999 CE syllabus had not changed. The ability of speaking, writing, and translation had been placed on the second level. All schools needed to meet the basic requirements of both levels to participate in the national CET-4 and CET-6 (Cai, 2015: 39). However, the CET Spoken English Tests were implemented in 1999 in some major cities in China (Jin & Yang, 2018), which had, to some extent, relieved the tension of the debate.

2.4.3 The Reform Period (2002-Now)

Teaching Method: diverse teaching methods

Moving to the millennium, the discussion of teaching method has been more

comprehensive and innovative, and it is agreed that CE teachers should tailor their teaching and adapt various teaching methods to their own contexts (Li, Xing & Wang, 2019). Classroom teaching has changed from teacher-centered to student-centered, focusing on the development of interests and autonomous learning. According to the specific situation of students, comprehensive teaching methods are adopted to explore and practice in the CE classroom. There are diverse teaching methods, such as task-based language teaching, cooperative language teaching, and experimental language teaching (Li, Xing & Wang, 2019). Task-based language teaching provides learners with interactive opportunities, and stimulates their ability to use language creatively to reflect the teaching concept of student-centered, autonomous learning and personalized learning (Zhang, 2010). The cooperative teaching focuses on the cooperative function between groups and individuals, improving the learning and teaching efficiency, optimizing the grouping according to the students' English achievements (Cai, 2019). The experiential teaching emphasizes the learning experience of students, so that students can lead the learning process and be responsible for the learning results. Specifically, in English teaching, teachers can optimize the teaching content by creating situations and ways to promote students' conscious participation in English learning activities (Zhu, 2022). In the reform period, the diverse teaching methods are more comprehensive and innovative. They focus on the student-centered concept, and the characteristics of college students in the new era, with the stimulation of students' interest in learning (Li, Xing & Wang, 2019).

Teaching Orientation: Emphasis for Listening and Speaking

Apart from teaching methods, teaching orientations have engaged in the debate between the emphasis on reading and the emphasis on listening and speaking. This time, priority has been given to listening and speaking. This debate was a continuation of the debate between English for literacy and English for practical use at the end of 20^{th} century.

In the spring of 2002, Zhang Yaoxue, director of the Department of Higher Education in the Ministry of Education, published an article entitled “Strengthening Practical English Teaching and Improving College Students’ Comprehensive English Ability” (Cai, 2015: 39). The main contents were as follows: (1) It criticized the current situation that many students’ listening and speaking had not improved with the phenomenon of “Dumb English”; (2) The goal of CE Teaching was mainly to stress reading, but it did not attach enough importance to the comprehensive use of English, such as listening and speaking; (3) It was proposed that listening and communication must be placed in an important position in English teaching (Cai, 2015).

However, some experts and specialists from CE had different views. They argued that reading for college students should never be neglected and denied (Cai, 2015). As expressed in their statements, “the current syllabus puts the reading first, which is a reasonable decision according to investigations and research, as well as the actual situation of our country. Under the current conditions in China, it is unrealistic to require all the college students to be proficient in reading, listening, speaking, writing, and translating. Although there is a higher demand for English speaking than in the past, for most college students, the main English language skills needed for future work are still reading” (Cai, 2015: 39-40).

Despite such opposition, the Ministry of Education had made a strong resolution to CE reform, this time, to shift focus to listening and speaking. First, the Ministry of Education had officially promulgated the College English Curriculum Teaching Requirements (the trial version in 2004). The teaching goal of CE was to teach students’ comprehensive English ability, especially listening and speaking, which requires colleges and universities to implement CE teaching with the new teaching orientation. Second, CET-4 and CET-6 had been reformed to include listening. Therefore, the proportion of listening rose from 20% in the past to 35% now (Cai, 2015). Three years later, the General Requirements for College English Teaching

(Ministry of Education, 2007) were issued with the teaching orientation emphasizing on listening and speaking:

> The aim of College English teaching is to teach students' comprehensive English, especially in listening and speaking, so that they can effectively communicate orally in English in their future work and social contacts.

2.5 COLLEGE ENGLISH CURRICULUM

The CE curriculum is further divided into 3 branches based on the General Guidelines of College English Teaching (Ministry of Education, 2020). For this study, the focus is on the General English and specifically English Listening and Speaking.

The General Guidelines of College English Teaching (Ministry of Education, 2020) were developed from the General Requirements for College English Teaching (Ministry of Education, 2007). In the General Guidelines, the nature of CE teaching is divided into two aspects: instrumental and humanistic. In alignment with these two aspects in the teaching nature, the three branches of CE curriculum are confirmed. The three levels of CE teaching are General English (GE), English for Specific Purposes (ESP) and Intercultural Communication courses (Wang & Wang, 2019). First, the General English courses are opened for college freshmen who are weak in English communication, which is necessary to improve general English ability through continuous learning (Wang & Wang, 2019). Second, the English for Specific Purposes (ESP) courses are oriented to the fields of professional English, aiming at enhancing students' ability to use English for professional and academic exchanges and work, and improving their professional and academic quality (Wang, 2016). Both GE and ESP courses emphasize the instrumental purpose of teaching. For humanistic purposes, the intercultural communication courses are set up in CE education, which can help college students to increase intercultural awareness and improve intercultural communication competence (Wang, 2016).

For this study, the context centers on GE courses, especially the course of English Listening and Speaking, as it is in alignment with the current teaching orientation, which emphasizes on the listening and practical communication of speaking in the General Requirements for College English Teaching (Ministry of Education, 2007). As this research study focus on speaking and spoken language production, the context of teaching English Speaking especially in the Reform Period of millennium China (2002-now) will be explored.

2.6 TEACHING ENGLISH SPEAKING IN MILLENNIUM CHINA

Moving toward the millennium, the teaching of English Speaking in China developed rapidly along with the booming growth of teaching reform in the CE Reform Period of China. In view of the shortcomings of CE Teaching, in the past, the current CE teaching has undergone reform with a focus on speaking and communication (Zheng, 2016). In addition to the traditional intensive reading and listening classes, many colleges and universities have offered oral English courses, such as listening and speaking, oral audio-visual, and oral English classes taught by foreign teachers, aiming at developing students' oral English (Zheng, 2016).

Over the past 20 years, the teaching reform of English Speaking in CE China has made rich achievements. Zheng (2016) has categorized the research scopes of the teaching reform in 3 levels: 1) the modes of teaching speaking; 2) problems and countermeasures in teaching speaking; 3) teaching methods of English speaking. As this study focuses on teaching methods, the reform of teaching methods will be examined.

2.6.1 Teaching Method Reforms in English Speaking

The teaching methods of English speaking in the CE Reform Period include

cooperative language teaching, experimental language teaching and task-based language teaching.

For cooperative language teaching, Zhu (2007) summarized four rules: quantity, quality, relevance, and manner, for applying cooperative principles to improving English speaking. Regarding the theory of experiential learning, Sheng (2010) designed teaching activities of questions and answers, cards, and group discussion. As for task-based language teaching, Zhang (2010) took a class of English speaking as an example to illustrate the application of TBLT, which indicated the teaching concept of student-centered, autonomous learning and personalized learning. Lan (2009) made a detailed introduction and description of the pre-task, during-task, and post-task design of teaching speaking, showing that TBLT could improve students' language production. Through the empirical research of teaching experiment and test analysis, Wang, Gao, and Cheng (2013) demonstrated the role of task-based speech in teaching speaking. Next, the specific teaching context of this study will be presented and the teaching context of English speaking will be analyzed.

2.7 TEACHING SPEAKING IN THE UNIVERSITY OF SCIENCE AND TECHNOLOGY

With respect to this study, the context focuses on teaching English Speaking as a General English course in a university of science and technology. This research is situated in XXX university where the author is currently working. Next, the teaching of English speaking in the university of science and technology will be investigated.

2.7.1 Challenges and Obstacles in Teaching Speaking in the University of Science and Technology

In the millennium China, opportunities and obstacles coexist in teaching English

speaking in the university of science and technology. As globalization and international communication has developed rapidly in the millennium China, the urgent needs of strong ability in English speaking are demanded in all walks of life. Therefore, the CE curriculum of English speaking has undergone serious reform. The improvement of English speaking has become the primary task of CE teaching in China. However, problems and obstacles still exist in the traditional teaching method as well as external and internal challenges in teaching speaking in the university of science and technology.

2.7.2 The Traditional Teaching Method in Teaching Speaking in the University of Science and Technology

Teacher-centered Classroom

The traditional teaching method, Grammar-Translation Method, of teaching speaking in the university of science and technology is often teacher-centered. The teaching contents focus, primarily, on form (Zou, 2014). While these could be helpful to develop students' reading, they are not beneficial for developing speaking (Liu, 2009; Zhang, 2014).

Stevick (1989) has stressed that "it is essential to put the learning principle (teaching from the perspective of students) before the teaching principle" (Li, 2007). Therefore, the success of oral English class largely depends on whether teachers and students are clear about their respective roles. First, students are the center and masters of the oral English class, and teachers are the organizers and planners. However, traditional teaching in the university of science and technology had often been the teacher-centered, spoon-fed process. The teachers were the center of the classroom, while the students passively acquired knowledge (Li, 2007).

In English speaking, language communication included two processes: the acquisition

of information and the language production. The traditional education of cramming grammar and vocabulary could not engage students to accomplish these two. Therefore, oral English class should be changed to student-centered activities and the teachers should step back as the organizers to guide the students to participate in various speaking practice so that they could actively participate in class (Li, 2007).

Focus on Form Practice

Apart from the teacher-centered classroom, the teaching contents of traditional speaking class usually focus on form, which is not the authentic real-life of English. Currently, even though many universities of science and technology attach importance to the practice of English speaking, most teachers of English speaking still adopt traditional teaching methods, which focus on form. In class, the speaking exercises are limited to reading passages, reciting, and practicing sentence patterns, and drill practice of conversation patterns (He, Li & Tao, 2017). The reason is that the teachers only pay attention to the form and contents of language and ignore the authentic use of language. This makes it difficult for students to achieve real-life communication and language production (He, Li & Tao, 2017).

2.7.3 Problems and Obstacles of Teaching Speaking in the University of Science and Technology

Apart from the traditional teaching method, the teaching of speaking in the university of science and technology still meets with external and internal challenges.

External Reason 1: Policy Orientations of National CE Development

Since the Reform and Open-up, the scale of teaching orientations tipped to English for literacy in the first 20 years. Then, in the last 20 years, emphasis shifted to listening and speaking. Given the fact that past traditions and influences still exist, the teaching of speaking meets with barriers and challenges.

ELT in China follows the general requirements and guidelines from the Ministry of Education. Taking the CE syllabus in China as an example, the focus was on the assessment of English reading and writing of college students for a long time before the millennium. It was not until 2004 that attention and focus were changed to listening and speaking. The university students nowadays majored in science and technology has been taught since primary school over the past decades (Zhang, 2014). In most of their English study history, emphasis was given to reading and writing, which would lead to the speaking of "dumb English". In addition, due to the long-term influence of traditional teaching methods, Chinese students have not paid enough attention to speaking and the comprehensive use of English (Zhang, 2014). Compared with English reading, English speaking is always in a weak position, especially in the university of science and technology (Zhang, 2014).

External Reason 2: Lack of Speaking Context

For students in the university of science and technology, the opportunity to speak both in and out of the English class is limited, which leads to the lack of practice and context in English speaking. From the perspective of linguistics, the speaking environment and context are of vital importance to the training of oral English (Zhang, 2014). In China, the environment of speaking English is deeply affected by various conditions. The occasions for students to practice oral English are mainly in the classroom. Apart from that, there are few speaking contexts for oral English practice. This situation seriously restricts students' enthusiasm in learning English (Zhang, 2014).

External Reason 3: Cultural Differences

Apart from the above external barriers and challenges, the culture and cognitive differences between China and the West could impede spoken language production for students in the university of science and technology. The culture differences between China and the West could lead to confusion and misunderstandings, which is

not accessible to students' language production. Culture is a complex concept, which includes arts, morality, beliefs, law, customs, and habits acquired by people as members of society. In communication, the choice of vocabulary could be a mirror, representing the national culture and customs (He, Li & Tao, 2017). Take the word "dog" as an example, meanings vary between China and the West. In Western countries, "dog" is usually given as a commendatory meaning, such as: "You are a lucky dog!" or "A good dog deserves a good bone.". However, in Chinese, "dog" is often given negative connotations of evil and brutality. Facing these, students in the university of science and technology can struggle with English learning, as well as English speaking (He, Li & Tao, 2017).

Besides, Chinese tend to speak more implicitly than people in the West. These differences could bewilder students in speaking and communication. Hall (2010) believed that high cultural context creates a high degree of dependence on the context, and many meanings are included in the context, while low cultural context has a low degree of dependence, and all information needs accurate speech expression. China belongs to a high context cultural society, which conveys information indirectly. Information often exists behind the language. However, many Western countries including English speaking countries of United Kingdom and United States are low cultural context. When communicating, they tend to be frank, direct to the themes, and often express information straightforwardly (Hall, 2010; He, Li & Tao, 2017). For learners of English speaking, the difference in thinking between China and the West could also hinder their understanding of speaking.

Internal Reason: Students' Own Problems

Considering the students themselves, most of them are introverted and silent in class, which is not active in English speaking. The character of Chinese students is bound to be marked with traditional Chinese culture: introverted, shy, and relatively silent in class (Cui, 2021). This has formed a strong contrast with European and American

students. In class, Chinese students' desire to actively participate in the oral practice has been seriously curbed (Zhang, 2014). Most students, generally, do not take the initiative to speak, let alone students in the university of science and technology. In a nutshell, the problem of students themselves is depicted as follows:

Inferiority Complex

In the university of science and technology, some students could form an inferiority complex, which could deter their practice of English speaking. Since the students in this context come from all over China, their oral English level is uneven, especially those from minorities and rural areas. Some of them are weak in language skills, which means they could distinguish the meaning of words, but cannot produce proper speaking. Still, some of them could produce simple vocabulary and grammatical structure, but they are weak in pronunciation and intonation (Liu, 2009). Therefore, they are afraid of making mistakes in class, so they have very low self-esteem. In addition, they are afraid of being criticized by teachers and ridiculed by classmates. Over time, they could form anxiety and inferiority complexes (Zhang, 2014).

2.7.4 Task-based Language Teaching (TBLT) in Speaking

Facing the challenges imposed by traditional teaching methods as well as the external and internal obstacles in teaching speaking, the adoption of tasked-based language teaching can address the problems with the following aspects: student-centered classrooms; focus on form and focus on meaning; governmental advocacy for TBLT; and TBLT enhancing cultural understanding and increasing learners' self-esteem in speaking.

Student-centered Classroom

For the theoretical foundation, TBLT is based on constructivism (Tian, 2013a), which holds the view that students are the main body of cognitive process and active builder

of knowledge. Therefore, teachers should change from knowledge imparter to helpers and promoters of students' active construction of meaningful knowledge. This requires teachers to change their roles to build the new student-centered classroom. Then, students can seize the initiative to actively participate in language learning and English speaking (Tian, 2013a).

To achieve the active participation in language classroom, TBLT has proved to be one of the best ways. According to Wei (2011), "task driven" teaching method is the "student-centered" teaching, breaking the traditional ways of language courses. The teaching system establishes the dominant position of students in teaching activities, encourages independent thinking, and fully mobilizes learning enthusiasm, initiative, and creativity to enhance active engagement in class.

Besides, it is further supported from the Report on the Quality of Higher Education in China, that task-based approach fits perfectly well with the requirements set by the quality of engineering education, which emphasizes student-centered approach for engineering students proposed by the International Engineering Union (Ministry of Education, 2016).

Focus on Form and Focus on Meaning

The TBLT teaching can be divided into focus on form and focus on meaning, particularly with the strong and weak form of task-based approach tailored to different teaching purposes in speaking. Willis (1996) has pointed out that the strong form of task-based learning activities focuses on the principle that meaning and practical use of the language are in priority in language learning. Likewise, Skehan (1998) has discussed tasks as follows: meaning and task completion come first. To summarize both viewpoints, the strong form of tasks should attach importance on how students communicate with each other, rather than what kind of language they use. However, the weak task type emphasizes that task is an important part of language teaching. It

advocates teaching language knowledge first, and then designs a series of tasks for students to practice doing things in language, and then summarizes language knowledge (Feng & Tang, 2004). In the speaking class for students in the university of science and technology, teachers can apply both types of tasks to suit their teaching purposes.

TBLT Advocated in the Government Document

In the recovery and development period (1978-2001) of CE teaching, as mentioned above, the traditional teaching method, Grammar-Translation Method, was adopted in English teaching. However, in the latest national document, the General Guidelines of College English Teaching (Ministry of Education, 2020: 33), task-based language teaching is promoted on the national level.

> College English classroom teaching can adopt task-based, cooperative, project-based, and inquiry-based teaching methods. The teaching concept should be changed to teachers as the leading role and students as the main body. The teaching should be characterized by teachers' guidance and inspiration and students' active participation in the language classroom.

According to the basic idea of constructivism, mentioned above, students' learning is essentially a process of actively constructing knowledge based on their own knowledge experience. The students' active participation and construction represent the state of high-quality classroom learning and teaching (Wang, 2016). Thus, effective instruction and teaching should be organized for students' learning (Tyler, 2014).

In the past, the actual situation of CE Teaching in China has shown that teachers' teaching is separated from students' learning with the traditional teaching method. The CE classroom teaching does not reflect the idea of "teachers as the leading role and students as the main body", lacking interaction and effective communication in the classroom (Wang, 2016). Therefore, in the General Guidelines of College English

Teaching (Ministry of Education, 2020), it is emphasized that teachers should improve the teaching methods and create active learning environment that is characterized by teachers' guidance and students' active participation (Wang, 2016).

TBLT Enhances Cultural Understanding

TBLT can enhance students' understanding for social customs and thinking differences between China and the West. For the social customs, teachers can use the task-based approach to start with English reading tasks on a regular basis. According to the specific contents of literary works or newspapers read by students, teachers can design topics that are close to their lives for open discussion (He, Li & Tao, 2017). Then, the students can be divided into several groups for discussion, so that each student can participate in the discussion. In this way, students can not only absorb and accumulate relevant cultural background and social customs in the process of reading, but also generate interest for oral communication (He, Li & Tao, 2017).

TBLT Increases Learners' Self-esteem in Speaking

As the student-centered teaching method, TBLT encourages meaningful teacher-student communication as well as the student-student communication, which promotes confidence in speaking.

In task-based teaching, students are viewed as the center of classroom. In the student-centered class, meaningful communication can be promoted through the communicative desire and activities between students and teachers (Lan, 2009). Meanwhile, group activities can strengthen cooperative learning among members. Additionally, the task can also create a relaxed and pressure-free language environment for students, so that they can communicate freely in English with both their teachers and classmates. Therefore, the fun of English learning is experienced along with the boost of self-esteem to form strong learning motivation (Lan, 2009).

2.8 AN INTRODUCTION TO THE UNIVERSITY AND THE COURSE OF ENGLISH LISTENING AND SPEAKING

Next, focusing on the specific research context, the university, the target students and the course of English Listening and Speaking will be introduced.

2.8.1 An Introduction of XXX University and the Target Students

The university under study is XXX University, which focuses on full-time undergraduate and postgraduate education. The undergraduate students in XXX university are students from all over China whose scores on the College Entrance Examination are over the line for the first-batch university. Most of the first-year undergraduate students can achieve very high scores in the College Entrance Exam. However, their English level of speaking is quite uneven, which has posed real challenges for English teachers. The freshmen of XXX university will take 4 credits of required English courses in their college education, which include 2 credits each for English Reading and Writing and English Listening and Speaking.

2.9 AN INTRODUCTION OF THE COURSE ENGLISH LISTENING AND SPEAKING

2.9.1 Course Objectives

For this research study, English Listening and Speaking is chosen as a required General English course for first-year college students in XXX university. Adopting the task-based approach, this course is designed to help students to promote listening and speaking skills as well as critical thinking and problem-solving skills. The objectives of this course are described as follows.

English listening and speaking, is a General English course for undergraduates studying in XXX university. This course provides students with knowledge and skills training of English listening and speaking by carrying out English audio-visual, oral training, and student-centered classroom activities. This course follows task-based teaching concept of "teacher-led and student-centered". It prepares a variety of language tasks for students in the classroom and provides rich, diversified teaching guidance for students. Through the study of this course, students will develop English listening and speaking skills and intercultural awareness. Additionally, students' critical thinking and problem-solving skills can be improved from the practice of real-life communication tasks.

2.9.2 The Course Structure

Based on the above course objectives, the course structure and contents are designed as follows: this course lasts for 16 weeks and every two weeks will cover one unit in the textbook, *New Horizon College English Listening and Speaking*. The specific contents can be seen in Table 2.2. Next, the teaching method, Task-based Language Teaching, with the focus on speaking will be presented.

Week	Contents
1-2	Unit 1 Traces of the past
3-4	Unit 2 A break for fun
5-6	Unit 3 Life moments
7-8	Unit 4 Getting from A to B
9-10	Unit 5 Relax and explore
11-12	Unit 6 Wit and fit
13-14	Unit 7 Weird, wild, and wonderful
15-16	Unit 8 Money matters

Table 2.2: Course Breakdown by Sections

2.9.3 Task-based Language Teaching in the Course for English Listening and Speaking

Task-based Language Teaching is applied in the course for English Listening and Speaking and it has the following characteristics: tasks as the main line, teachers as the leading roles and students as the center. The teachers play the role of guidance and inspiration and provide opportunities for speaking practice. Through cooperation and interaction, students can be guided to think actively and practice creatively. Through the explanation of new knowledge, cultural points, and speaking skills, students can take the initiative in speaking, so that the teaching activities can truly realize the transformation from “teaching” to “learning”.

Task Types for Speaking

Next, the task types for speaking will be analyzed from the perspective of interaction and communication as well as the viewpoint of information processing.

Focusing on the interactive and communicative perspective in the speaking tasks, Prabhu (1987) put forward three types of tasks: information-gap tasks, reasoning-gap tasks, and opinion-gap tasks. To further develop the tasks for interaction and communication, seven types of tasks were raised by Pattison (1987), which were questions and answers, dialogue and role play, matching activities, communication strategies, pictures and picture stories, puzzle and problems, discussion, and decisions.

In this course, English Listening and Speaking, there are a variety of interactive speaking tasks, such as conversation practice and speech-making. During this class, students can practice conversations with vocabulary, sentence patterns, and expressions in various situations. For the speech-making, the teacher will provide students with the topic and outline, and at the same time, provide guidance to operate the whole task.

Aside from the interactive tasks, Nunan (2011) depicted one-way task and two-way tasks for speaking from the interpretation of information process by Bygate (1987). According to Bygate (1987), the ways a speaker processes and expresses the information came in two dimensions: one for providing information; and the other for social interaction. The layer of social interaction was described in Prabhu (1987) and Pattison (1987)'s task lists for oral speaking, whereas the layers for providing information involved interpreting and evaluating information (Bygate, 1987). For interpreting and evaluating information, the speaker would describe, contrast, or justify his/her opinions (Bygate, 1987). Related with Bygate's (1987) viewpoint, Nunan (1989) generated the monologic task, a one-way expression task related to the expressive competence of language, in contrast with the dialogic task. For the monologic, one-way task, summarizing, retelling, and impromptu speeches are most common (Yu & Lei, 2004; Ou & Huang, 2016).

For this speaking course, speech-making tasks are designed to enhance critical thinking. After discussion and cooperation, each student can develop his/her own ideas in the practice. Therefore, the speech-making tasks will receive focus. Next, the speech-making tasks will be introduced.

Speech-making

Speech-making is common in social practice, which is essential for students' study and life. Practicing the speech-making task is beneficial for students, not only for enhancing spoken language production, but also promoting independent and critical thinking, as well as boosting self-esteem (Yan & Jia, 2006).

Speech is a kind of social practice activity. It is the behavior that the speaker is at the center of the communication process to provide information and present ideas to the audience. This is a relatively advanced language communication activity, and it is also an indispensable part of human activities (Yan & Jia, 2006). Presenting a speech

makes people understand one's thoughts and opinions. For example, when the students elaborate views on certain topics in class, or when they compete for employment, conduct business negotiations, self-promotion, introduce products, and communicate ideas in future work, all the above requires speech-making skills.

In the CE speaking class, the practice of speech-making can develop students' proper English expression ability, encourage independent and critical thinking. Meanwhile, it is helpful to enhance students' courage, confidence, and immediate reaction (Yan & Jia, 2006).

2.10 CHALLENGES AND RESEARCH QUESTION

However, the design and implementation of speech-making tasks in the speaking classroom remains challenging for CE teachers.

For the design of task characteristics, the familiarity with speech topics and the structure of speech-making tasks (Skehan, 2011) are unclear for language teachers. For instance, teachers are unsure which topic "shared-bikes' influence on our life" or "describe a trip in Shenzhen" is more familiar to students. How can the varied familiarity of topic and structure impact students' spoken language production? Meanwhile, in the process of task implementation, the task conditions of pre-task planning (Ellis, 2009) (strategic planning: brainstorming for ideas, providing model speeches; rehearsal) are uncertain for language teachers on their influence of students' spoken language production.

Still, teachers of English speaking have little idea as to the pedagogical interventions in task design and implementation to facilitate students' spoken language production, because there are few research findings available, especially in the context of university of science and technology in China.

Therefore, this study will investigate the following research question:

How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?

2.11 CONCLUSION

Overall, this chapter has presented the context of this study. The scope of the research was narrowed from English as a global language, ELT in China to CE in China. Specifically, the research study is set within the speaking course in a university of science and technology in China. Considering the problems and challenges which have occurred in teaching speaking, TBLT was selected. In the speaking course, one type of the task, speech-making, was chosen for further study. When designing and implementing the speech-making tasks in class, the task characteristics, task conditions, and their impacts on students' spoken language production, which presently remain unknown to the teachers, should receive careful consideration. In the next chapter, the themes of this study will be elaborated in the literature review.

Chapter 3 Literature Review

In order to answer the research question in previous chapters: "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", the literature review section will probe three themes of tasks; spoken language production; task and spoken language production with the influence of task characteristics and task conditions on spoken language production.

Research Question	Literature Review	
RQ: How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?	A. Tasks	1. Conceptions of TBLT 2. Definitions of tasks 3. Task types and speech-making genres
	B. Spoken language production (CAF)	1. Levelt's (1989) model of first language speech production 2. Second Language Acquisition 3. Spoken language production constructs: Complexity, Accuracy and Fluency (CAF) 4. Psycholinguistics information process 5. VanPatten's (2007) Input Process Model 6. Cognitive factors on second language production 7. Skehan's (2014) Limited Attentional Capacity Model
	C. Tasks and spoken language production	1. Task characteristics and spoken language production 1) Familiarity of information 2) Degree of structure
		2. Task conditions and spoken language production 1) Strategic planning and spoken language production 2) Rehearsal and spoken language production
	D. Research gap	
	E. Conceptual framework	

Table 3.1: The Literature Review

The tasks, conceptions, definitions, types, and specifically the speech-making genres

will be explained. Additionally, the complexity, accuracy and fluency constructs of spoken language production will be established. To link the two themes: tasks and spoken language production together, the task characteristics and task conditions will be reviewed in the design and implementation process. Finally, the impacts of task characteristics and task conditions on spoken language production will be demonstrated along with pedagogical implications for the speech-making tasks. As the design and implementation of speech-making tasks and their impacts on spoken language production cannot be found in previous studies in the designated context, the research gap is shown to justify the design of this study. To sum up this section, a conceptual framework will be demonstrated. For the clarification of this chapter, the literature review is divided into the above sections (Table 3.1)

3.1 TASKS

To scrutinize the theme tasks, the conceptions of Task-based Language Teaching (TBLT) and the definitions of tasks will be displayed. Then, the task types will be introduced and the speech-making genres will be concentrated.

3.1.1. Conceptions of TBLT

TBLT has originated from Communicative Language Teaching (CLT), and is considered as a strong representation of it. TBLT is learner-centered and learning through doing tasks, which shows the equal relationships between teachers and students (Adamson, 2004b, Long, 2015) (see Figure 3.1).

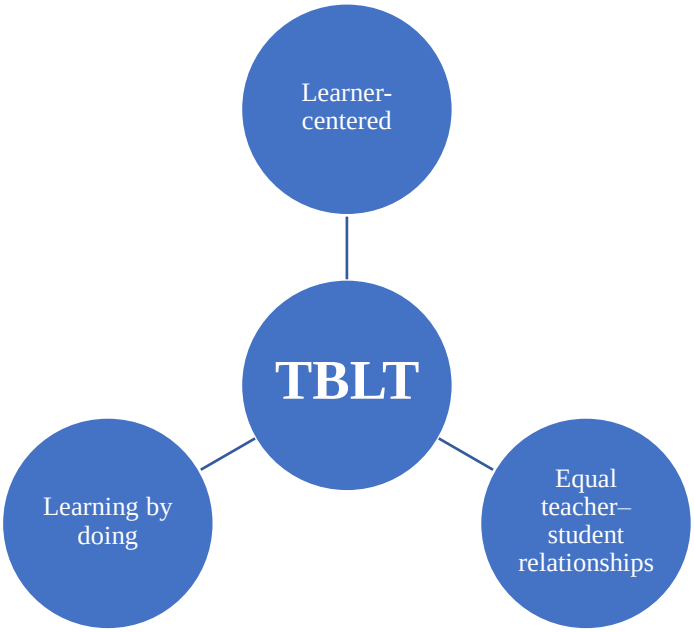


Figure 3.1: Conceptions of TBLT (Long, 2015)

English Language Teaching Methodology

Towards the 1960s and 1970s, ELT was associated with psychology and authentic, real-life settings. Depending on the problem-solving skills from learners, the Silent Way encourages learners to adopt psychological inductive and self-monitoring techniques for the construction of structural knowledge in the target language (Adamson, 2004b). For the realistic perspectives, the Total Physical Response method fosters learning through reacting to the physical instructions by the learners. Another psychological and real-life example is called Suggestopedia, generating from the real-life settings of yoga and psychotherapy. For Suggestopedia, a comfortable and relaxing environment with music is provided for the alleviation of stress during learning (Adamson, 2004b).

In the 1970s, CLT in ELT became most popular, which could be traced from two origins. One is like the psychological and authentic aspects from Suggestopedia. Having developed the ideas from Charles Curran, CLT aims to relieve learners'

pressure by simulating the techniques in the psychological counseling process (Knight, 2001). The other is to view language as a social practice and to express meaning, which generates communicative competences through social interaction and communication (Adamson, 2004b; Zhou, 2019).

Under the umbrella of CLT, there exists the weak and strong models. The Functional-Notional Approach was associated with one of the weaker forms. The main features are to emphasize the functions of language, such as inquiry and order, as well as the notions of language, such as time and space (Adamson, 2004b). To complete the teaching process of Functional-Notional Approach, the students will read and recite words, phrases, sentences, paragraphs, and practice based on mechanical imitation and learning (Wang & Wang, 2008). Even though it is used in the realistic context for communication, the Functional-Notional Approach aims at the targets of linguistic items, which is regarded as the weak form of CLT (Adamson, 2004b).

In contrast, TBLT is considered as a strong form of CLT. By the 1980s, researchers and teachers had paid more attention to TBLT (Adamson, 2004b; Wu & Pan, 2012). The earliest implementation and experimentation of the task-based approach was the Bangalore projects by Prabhu (1983). As a teaching approach initially belongs to CLT, TBLT is a learner-centered approach, based on the use of real-life tasks as the core unit of planning and instruction in language teaching (Richards & Rogers, 2001). Along with the authentic and communicative tasks, TBLT was developed from the concept of *l'educacion integrale*, which is closely associated with the concepts "learning by doing" to stimulate learners' interest and internal motivation (Fang, 2003). Besides, TBLT emphasizes equal relationships between teachers and students, establishing beneficial psycholinguistic conditions for language learning (Long, 2015). Overall, the task-based approach contains the ideas of communicative approach, such as communication-based thinking, but it develops and transcends the communicative approach (Chen, 2008).

TBLT: Learner-centeredness as the Core

Based on the philosophical psychology of "constructivism", TBLT reflects the principle of learner-centered concepts (Candlin, 1987; Fang, 2003). Constructivism holds the view that knowledge is constructed by learners and influenced by social culture. This theory believes that human cognition develops together with experience and knowledge generated from the reorganization and reconstruction of experience, discovery, and creation (Fang, 2003).

For this study, the task characteristics and task conditions and their influences on spoken language production will be explored. For spoken language production, the formation of language is a cognitive construction process, which can be understood as the psychological process of meaning construction (Wang, 2010). In this study, the task characteristics and task conditions can be varied to influence the psychological construction process of spoken language production.

For learner-centeredness, the task chosen in this study, speech-making, could provide learners with real and meaningful goals to produce spoken language during the psychological construction process. With the meaningful purposes, it is bound to activate learners' existing knowledge and experience and inevitably stimulate learners' desire for new knowledge (Fang, 2003). Such desire for knowledge from learners has been elaborated by Long (2015) for the following two reasons: First, initiated by the students' current or future needs of communication, the contents of tasks can be determined according to their needs of language production. Second, the linguistics knowledge learned in the speech-making tasks can foster students' keen willingness to learn and speak (Long, 2015).

Therefore, it is through the participation of speech-making tasks that the learners show their readiness to learn, promoting the reorganization and reconstruction of knowledge. Moreover, once the learners experience the significance of participation

and success of new knowledge construction in speech-making, they will be more active and motivated, fostering the internalization of language learning (Fang, 2003). That is why the core of the task-based teaching method is “learner-centered”.

TBLT: Learning by Doing the Real-world and Meaningful Tasks

The learner-centered TBLT approach emphasizes that language learning is acquired through doing real-world and meaningful tasks (Samuda & Madden, 1984; Zeng, 2005).

As mentioned above, the concept of *l’educacion integrale*, in conjunction with the concept of “learning by doing”, was first expressed by the French utopian socialist Fourier, which meant the whole person’s education would be combined with both mental and physical learning (Long, 2015). The positive value of learning by doing has integrated intellectual reasoning abilities and practical experiences together. Through people’s personal experiences and practical work with real-world tasks, learning can be accomplished with abstract concepts and knowledge in daily life (Long, 2015).

The best part of TBLT is that through students’ doing the tasks by themselves, the personal experience of such practice can ensure the mental learning of language. Meanwhile, the students’ internal drive will be maximized and their ability will be developed to think about and solve the problems. During the learning process in this study, speech-making tasks can provide students with the opportunity to use language while doing the tasks for communication, coordination, and cooperation with others (Xia, 2001). Through completing the tasks, students can internalize the learning process and promote language production (Feng & Tang, 2004).

TBLT: Equal Teacher-student Relationships

Unlike traditional teacher-centered method, TBLT encourages equal relationships

between teachers and students, which can establish the learners' active roles in learning and effectively mobilize positive emotions and motivation.

In traditional teaching, the hierarchical, authoritarian class can often oppress students' needs in language learning. In contrast, egalitarian teacher-student relationships have been acknowledged and recommended to promote learning (Long, 2015). In TBLT classroom, the equal teacher-student relationships can create an active learning environment of students through the role-switching between teachers and students. When the teachers' roles are to support students' interests, the students will take the initiative in learning. When the learners have taken active control of their studies, their learning will be motivated (Nunan, 2011). In these cases, the teachers should be adaptive as the facilitators, participants, or observers in the learning process. Finally, the egalitarian teacher-student relationships can not only provide a positive classroom climate for language learning, but also promotes student's active roles in learning to stimulate interest and motivation (Long, 2015).

However, counter-arguments have been found in the Chinese context when discrepancies have existed between the expectations of the roles between teachers and students. In the TBLT classroom, the teachers can be the guiders, task designers, motivation stimulators, and monitors (Su, 2012). However, empirical research has indicated that some teachers do not assume a good role of task designers (Wu & Pan, 2012). For the class motivation, some students were still accustomed to the traditional classroom setting. They were not willing to participate in TBLT. Furthermore, some other students were reluctant to be monitored by the teachers in the classroom (Shen, 2013; Zheng, 2010). In these cases, the egalitarian teacher-student relationship cannot be guaranteed.

To conclude, obstacles do stand between the egalitarian ideals and real-world practice. Nevertheless, the teachers should enhance communication with the students to

understand their practical needs, help them to understand their own roles in the learning process, and have constant reflections on classroom practice (Wu & Pan, 2012; Zheng, 2010). Therefore, encouraging the equal relationships between teachers and students and overcoming the challenges in classroom practice can ensure the beneficial effects of TBLT.

Overall, the concepts of TBLT from its origins in ELT methodology have been explained. Moreover, TBLT is analyzed from the 3 perspectives of learner-centeredness, learning by doing and equal teacher-student relationships. Next, the definitions of tasks and task focus will be investigated.

3.1.2 The Definitions of Tasks and Task Focus

The Definitions of Tasks

In order to better understand TBLT, the definitions of tasks from researchers and scholars are reviewed in the following table (see Table 3.2). From the early researchers and scholars in TBLT, the tasks are considered as the activities that encourage language learning with an outcome (Crookes, 1986; Prabhu, 1987; Richards et al., 1985). Later, as a strong model of CLT, the meaningful communicative process and outcome are emphasized through the definitions of tasks (Breen, 1989; Bygate, Skehan & Swain, 2001; Candlin, 1987; Lee, 2000; Willis, 1996). Then, the tasks are further categorized as real-world tasks and pedagogical tasks depending on their focus. Real-world tasks focus on meaningful communication, while the pedagogical tasks focus on the linguistic forms (Nunan, 2011; Long, 2015).

After reviewing the above definitions, Ellis's (2013) definition has been selected to provide a basis for the study. A task is a differentiated, goal-oriented process, with a number of steps, which follows a series of cognitive and communicative procedures, and has a realistic communicative outcome. Additionally, a task is sequential and can

be subject to pedagogical intervention.

Authors	Task Definitions
Richards et al (1985)	Tasks are the **activity** that people carry out after learning, understanding, and experiencing language.
Crookes (1986)	Tasks are the work or **activities** with a clear **purpose.**
Prabhu (1987: 24)	(Tasks) are the **activities** which required learners to arrive at an **outcome** through some process of **thought,** and which allowed teachers to **control and regulate** that process.
Candlin (1987)	(Tasks are) a set of differentiated, sequencable, problem-posing **activities** involving learners and teachers in some joint selections from a range of varied **cognitive and communicative** procedures in the collective exploration and pursuance of foreseen or emergent **goals** within a social milieu.
Breen (1989)	Tasks are "the structured **plans** for the provision of opportunities entailed in a new language and its use during **communication**".
Willis (1996: 23-25)	Tasks are always the **activities** where the target language is used by learners for a **communicative purpose** to achieve an **outcome**.
Lee (2000)	The task is (1) a classroom **activity** or exercise, including (a) a **goal** that can only be achieved through **communication** between participants; (b) the natural process of organizing and arranging communication; (c) paying attention to meaningful communication; (2) an attempt to understand, practice and use the target language when a learner implements a learning **plan.**
Bygate, Skehan & Swain (2001)	Tasks refer to the **activities** that require learners to use language to achieve their **goals** based on emphasizing language meaning.
Nunan (2011:4)	The tasks are divided into two types: **real world tasks and pedagogical tasks.** Real world tasks are used in the places out of the classrooms. Pedagogical tasks are "pieces of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on grammatical knowledge to express meaning, and in which the intention is to **convey meaning** rather than manipulating form".
Long (2015)	TBLT (upper case): real world **activities** when people plan, conduct, and recall their days; tblt (lower case): classroom tasks in commercially published textbook that are focused on **forms**.
Ellis (2013: 2)	6 dimensions should be considered: **scope, perspective, authenticity, linguistic skills, psychological processes, and outcomes.** Tasks should be viewed based on their scopes such as workplans, perspectives focusing on meaning, authenticity of real-world processes, or any of the 4 linguistic skills, the cognitive processes, and the communicative outcomes.

Table 3.2: A List of Task Definitions by Researchers and Scholars

Tasks Focus

After the clarification of task definitions, the tasks which focus on form and focus on

meaning will be distinguished. Then, the continuum (see Figure 3.2) used by Tong et al (2000) depicts the weak, medium, and the strong types of tasks, with the form focus on the weak end and meaning focus on the strong end. For this research study, the strong types of tasks with purposeful, authentic, and open-ended outcomes will be chosen.

Focus on Form versus Focus on Meaning

Long (1991: 44) has explained that "focus on form" stresses "the content of syllabus and lessons based on the linguistic items themselves include (structures, notions, lexical items, etc.)". Long and Robinson (1998) have further illustrated that "focus on form" refers to the basis of the form of language in traditional language teaching. The class design on the structure of language is the main organizational principle, and the simple accumulation of isolated language phenomena is the main purpose. Conversely, "focus on meaning" requires learners to concentrate on the meaning and messages to express (Zhang, 2006).

To link the two types of instruction "focus on form" and "focus on meaning" with task, Long (2015) has distinguished the difference of "task-based language teaching" (lower case) and "Task-Based Language Teaching" (upper case) as mentioned above. "tblt" in the lower case is an activity or exercise carried out in a classroom situation with its syllabus designed to grammatical structures and vocabulary. However, TBLT in the upper case, focuses on meaning and refers to an activity using authentic texts as the basis for completion.

By taking a more comprehensive view of the use of tasks, Littlewood (1992) has regarded the tasks with the focus on form and meaning as a continuum with focus on language as a medium, on one end, and focus on learners' message on the other. Drawing upon Littlewood's continuum and the weak, medium, and strong categories of tasks from Legutke and Thomas (1991), Tong et al (2000) have produced a

framework for the types of tasks in ELT (Figure 3.2).

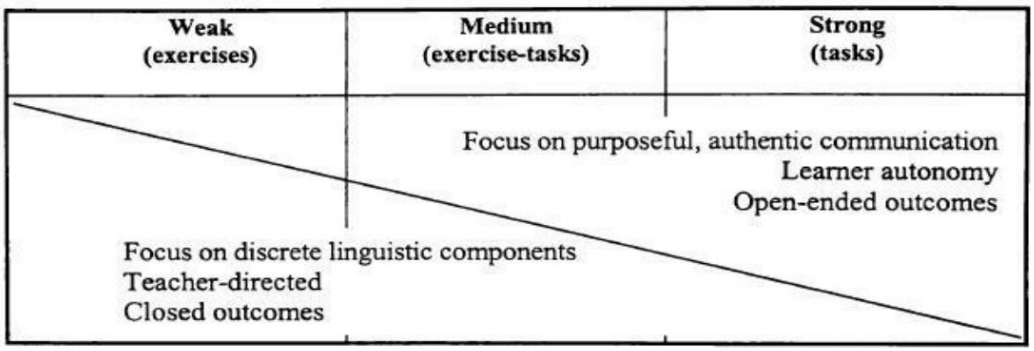


Figure 3.2: Types of Tasks (adapted from Tong et al, 2000)

The nature of English Speaking in CE is to foster students' spoken language production, instead of emphasizing language forms, so the strong types of tasks with purposeful, authentic communication, and open-ended outcomes will be chosen for this research. Next, the categories of task types will be investigated.

3.1.3 Task Types

The various task types can be distinguished in accordance with the different classifications. As mentioned in Chapter 2, the tasks can be categorized from the angles of interaction and communication, as well as information processing. However, there is the overlapping of tasks from both categorizations (Table 3.3).

Focusing on the information processing of interpretation and evaluation, the task selected in this study is the monologic, one-way task of speech-making, in which speakers can describe, contrast, and justify their opinions (Bygate, 1987). In this case, learners' language competence and critical thinking can be developed. To further understand the speech-making tasks, the different genres of speech-making will be identified.

Task Classifications	Task Types
Interaction and communication	Information-gap tasks, reasoning-gap tasks, and opinion-gap tasks (Prabhu, 1987)
	Questions and answers, dialogue and role play, matching activities, communication strategies, pictures and picture stories, puzzles and problems, discussion, and decisions (Pattison, 1987)
Information processing	One-way tasks and two-way tasks (Nunan, 2011) Monologic tasks and dialogic tasks (Nunan, 1989)

Table 3.3: Task Classifications and Task Types

3.1.4 Speech-making Tasks

For the speech-making tasks, there are various genres based on how the speakers deliver the speeches. Among the speech genres, there are readings with a manuscript, recitation speeches, outlining speeches, and impromptu speeches (Bao, 2019).

Reading from a Manuscript

Reading from a manuscript refers to the fact that the speaker has prepared a manuscript for the contents before the speech. During the speech, the speakers will read it to the public with reference to the written manuscript (Bao, 2019). This kind of speech is usually used for serious occasions, such as official meetings or reports. The language structure and expressions of this kind of speech have been prepared in advance, which emphasizes accuracy and preciseness (Bao, 2019). However, in this type of speech, the speaker maintains less interaction and emotional communication with the audience, and the acceptance of the speech from the audiences is poor (Bao, 2019).

Recitation Speech

For the recitation speech, the speaker will also prepare the written manuscript before the speech. Unlike reading from the manuscript, the speaker memorizes and recites the whole text of manuscript before the speech (Bao, 2019). Although the recitation speech does not use the method of reading the manuscript, the speaker's thinking is

largely dependent and limited in the contents and structure of the speech, which cannot guarantee authentic communication between the speaker and audience (Bao, 2019).

The Outlining Speech

For the outlining speech, the speaker will take notes on the structure and content of the speech with an outline. When the speaker expresses his own views, he will expand his thinking according to the outline and organize his language on the spot (Bao, 2019). In this way, the speaker will have high flexibility and ample room to develop his ideas, combining the emotional expressions with the audience to achieve interaction and communication. This type of speech not only has good organization and high appeal, but also can be a training method to improve the level of speech for beginners (Bao, 2019).

The Impromptu Speech

The impromptu speech refers to the speech in which the speaker has no prior preparation or prompt in advance. The speaker can express his/her opinions on the topic (Bao, 2019). This kind of speech is very difficult but provides room for the speaker to develop immediate reactions, speaking skills, logical, and critical thinking. When making the impromptu speech, the speaker can fully grasp the initiative of speaking and combine the reaction of the audience to practice social interaction skills (Bao, 2019). Meanwhile, it will require the speaker to have good memory, adequate use of language and vocabulary as well as the courage to speak logically and critically in public. Therefore, the impromptu speech is a good way for college students to master speaking skills and develop comprehensive abilities (Bao, 2019).

For the speaking class in this study, both reading from the manuscript and reciting the speech will not be effective as they could not generate meaningful and authentic communication between the speakers and the audiences. Whereas, both the outlining

speech and the impromptu speech will be chosen in this research as they can pass the speakers' viewpoints and information as well as developing students' spoken language production in the meaningful and purposeful communication process. At the beginning of the speech-making tasks, teachers can set up brainstorming, exchanging ideas and negotiating meanings among the learners to ensure the meaningful and purposeful communication. When the speeches are done in class, the students can hold discussions with the speakers to exchange ideas of their viewpoints. Accordingly, the strong types of speech tasks focusing on meaning are designed in this study.

The Speech Genres

The outlining speech and the impromptu speech require the learners to express their views on the scene and situation without preparation. These improvised speeches have seven spoken style features (first and second person pronouns, vague words, emphasis particles, abbreviations, verbs, adverbs, conjunctions) and seven written style features (passive voice, nouns, prepositions, adjectives, determiners, lexical complexity, lexical diversity) (Yu, Lu & Sun, 2010). Although the outlining and impromptu speech styles belong to the oral output, they are different from daily conversations, and have their own characteristics of both oral and written languages. Therefore, in general, both the outlining speech and the impromptu speech style exist with the mixing oral and written styles (Yu, Lu & Sun, 2010).

Overall, in the first theme "tasks", the conceptions, definitions, and types of tasks have been discussed. Particularly, the speech-making genres and the selected outlining speech and the impromptu speech are explained. Moving to the next theme, "spoken language production" will be reviewed.

3.2 SPOKEN LANGUAGE PRODUCTION

To examine the theme "spoken language production", a review will be conducted of

the first language speech production and Levelt's (1989) Speech Production Model to view how language is produced. Linked with this research context, the second language production will be analyzed with second language acquisition and the spoken language constructs: Complexity, Accuracy, and Fluency (CAF). Then, from the perspectives of psycholinguistics, the information process, and VanPatten's (2007) Input Process Model in second language acquisition will be analyzed. After that, the cognitive factors that would influence spoken language production will be indicated. Finally, Skehan's (2014) assumption, the Limited Attentional Capacity Model, will be presented.

3.2.1 First Language (L1) Speech Production

Language production is a psychological process in which language expresses ideas. By means of encoding the ideas, language and psychology, the sounds and words of a certain meaning are sent out by means of the articulator (Zhang, 2019).

Such psychological processes of spoken language production are regarded as a complex and multi-faceted production, which is explained by psycholinguistics through models ever since the 1970s (Liu, 2020a). In the fields of psycholinguistics, there are three influential and recognized models of L1 speech production: Fromkin's Serial Model, Dell's Parallel Model, and Levelt's Modular Speech Production Model (Liu, 2020a).

Fromkin (1973) has adopted a series of speech errors to observe the roles of language units in speech production. Eight types of speech errors were summarized: transfer, exchange, advance, delay, increase, decrease, substitution, and combination. However, Fromkin's speech error analysis has indicated that speech production goes through many independent stages, which can make all levels of language production isolated. Furthermore, not enough evidence for speech errors in interpreting the unprepared

speech was found (Liu, 2020a).

After Fromkin's serial model, Dell (1986) has proposed the parallel model, which is interactively activated among semantics, syntax, morphology, and phonetics levels. Dell's parallel model has assumed that all levels are operated at the same time (Liu, 2020a). Nevertheless, problems have existed in whether the activation of phonemic information in morphemes is simultaneous (Zhou, Zhuang & Shu, 2001). In order to prevent the simultaneous activation of phonemes between different syllables causing errors, Dell has proposed the so-called "binding by time" mechanism, which considers that phonemes are combined in turn with morphemes or syllables. Even so, experimental studies have shown that when speakers have multiple speech forms at the same time, they do not integrate the speech information, which results in errors (Zhou, Zhuang & Shu, 2001).

The last one, Levelt's (1989) modular speech production model, is based on decades of psycholinguistic research and many empirical studies, which is the most influential model of L1 speech production and the most widely used theoretical framework of speech production (Liu, 2020a). Therefore, Levelt's model will be chosen to view the language production process and later lead to the assumptions of Skehan's Limited Attentional Capacity model in this study.

Levelt's (1989) Model of First Language (L1) Speech Production

Levelt (1989) has proposed the information processing model of L1 speech production in three hierarchically modular stages: conceptualization, formulation, and articulation. The conceptualization stage is to develop and organize the ideas into a communicative goal. Then, in the formulation stage, a phonetic plan is made for the content of speaking. Finally, articulation is created when the phonetic plan is transformed into the actual speech (Ellis, 2013). To scrutinize the L1 speech production process, the following figure can be referred to (Figure 3.3). In this figure,

the L1 speech production involves several stages. For each stage, there will be some input and output. The output of one stage could be the input in another (Levelt, 2008). Based on the map in Figure 3.3, the 3 stages of this information process will be explicated.

Conceptualization

The first stage is conceptualization. In this stage, the speakers need to generate mental activities involving the conceptualization of speaking intentions before speaking. For such an intention, the speaker would need to decide what information is necessary to the speech, how the information should be ordered, and follow up with the expressions to realize the intention (Levelt, 2008). The mental activities of conceiving, planning, selecting, and monitoring are considered as the process of conceptualization. The output of this stage constitutes the preverbal message (Levelt, 2008).

Formulation

Moving to the formulation stage, the preverbal conceptualized messages are accepted as input. Then, those messages are transformed into the linguistic messages, which consist of two parts (Levelt, 2008). The first is grammatical encoding including the procedures of accessing lemmas (the basic form of a word) and the procedures of all the syntactic building (the way that words and phrases are put together to form sentences). Then, a surface structure, "an ordered string of lemmas grouped in phrases and subphrases of various kinds", will be produced by the grammatical encoder (Levelt, 2008: 11). Based on the surface structure, the second part of the linguistic message, the phonological encoding is established with the plan of building the phonetic or articulatory utterance for the lemmas and forms. The output of the phonological encoding comes with a phonetic or articulatory plan (Levelt, 2008).

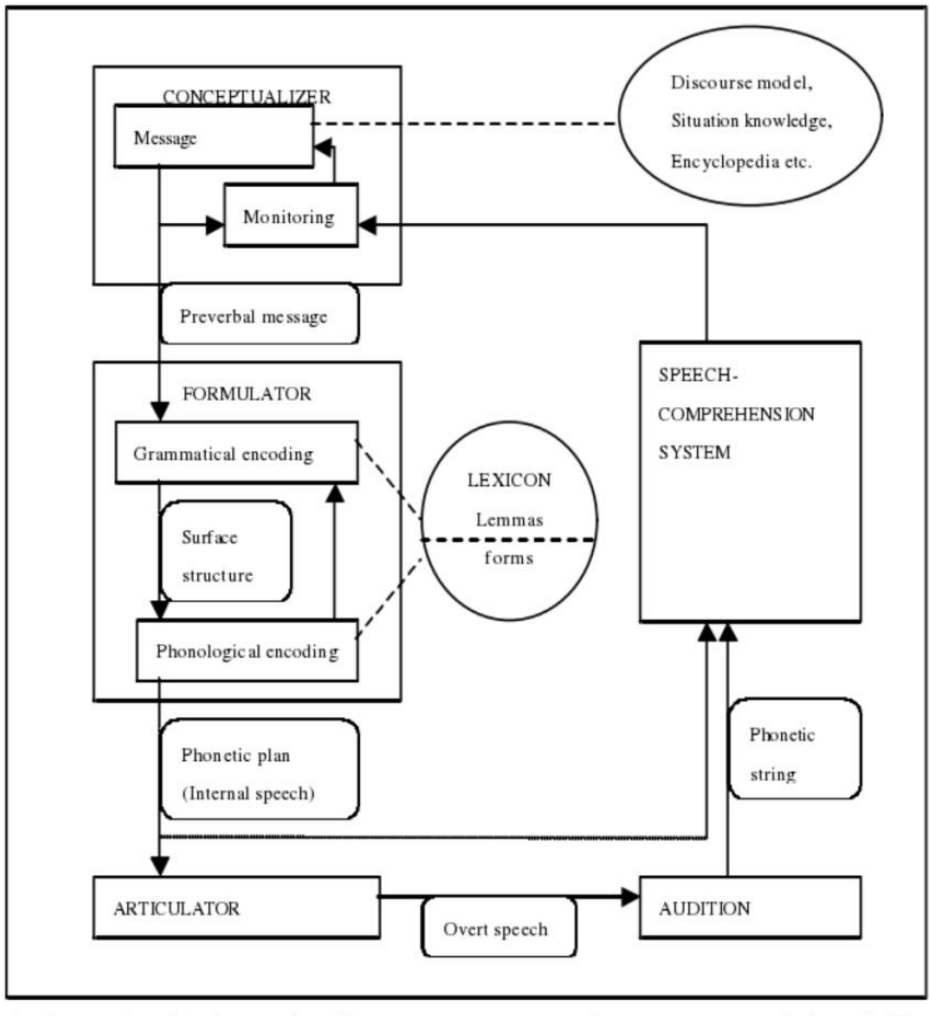


Figure 3.3: A Map for the First Language Speech Production Process

Articulation

The articulation stage, the end product, which is the phonetic or articulatory plan of the formulating stage, turns into the input of this stage. The phonetic plan is realized and delivered. In this execution stage, the internal speech of phonetic plan can be generated faster than the articulation (Levelt, 2008). Then, there is an articulation buffer as the temporary storage of the phonetic plan. The speakers can reclaim the phonetic plan from the buffer and deliver them in the articulation stage. The output of the articulation is called overt speech (Levelt, 2008).

Based on the above modular speech production theory by Levelt (1989), Kormos

(2006) has further summarized the four elements in the process of speech production: conceptualization, formulation, articulation, and self-monitoring. Conceptualization is a process of preparing the contents of a speech. Then, the formulation includes the encoding of grammar, vocabulary, and pronunciation of the message. With the formulation ready, the articulation is the output of speech. After that, self-monitoring includes checking the accuracy and appropriateness of the output. In L1 speech production, both phases of conceptualization and self-monitoring require cognitive attention, while the formulation and articulation are natural and automatic for L1 speakers (Xu & Chen, 2018).

The above analysis concentrates on Levelt's model of L1 speech production. However, the context of this study focuses on second language speech production for students in a university of science and technology in China. Then, based on Levelt's model of the language production process, second language speech production will be scrutinized.

3.2.2 Second Language (L2) Speech Production

To fully understand the concepts of L2 production, first, the understanding of learners' second language acquisition (SLA) should be reviewed. In SLA, the scope and the origin of SLA research are analyzed. Then, the Input and Output Hypotheses are raised. Next, based on the Output Hypothesis, the triad constructs of complexity, accuracy, and fluency (CAF) are identified by Skehan (1999), which can serve as the measurements of L2 speech production.

Unlike L1 production, L2 speech production, however, can be more demanding in the cognitive information process and the memory system, along with insufficient mental lexicons, the storage of "considerable information about each lemma and information", to support the natural and immediate L1 speech production (Skehan, 2011: 253).

Therefore, in the process of L2 speech production, the information and memory system are cognitively challenging for L2 learners. The role of cognitive factors is more complex, because L2 speakers have different degrees of conceptualization and formulation (Levelt's model) for L2 language coding (Xu & Chen, 2018). The pre-verbal message in the conceptualization stage is difficult to produce in L2 context. Meanwhile, the L2 speech formulation requires a conscious attention search mechanism to extract the appropriate lemma, cooperating with activated concepts, to complete the syntactic and lexical coding process (Xu & Chen, 2018). With smaller, incomplete, less organized, and less redundantly structured mental lexicons, the formulation stage will be considerably difficult for L2 speakers to find ways and resources to express meanings (Skehan, 2011). Thus, the cognitive factors in L2 production are directly influential to the quality of L2 output (Xu & Chen, 2018). Consequently, the cognitive perspectives which could impact L2 production will be investigated after the clarification of the CAF constructs in this section.

The Scope of Second Language Acquisition (SLA)

The scope of SLA begins with fundamental inquiries of how SLA occurs, which in other words, is how an L2 learner blend in the internalization of L2 linguistic system (VanPatten, 2007). To a broader extent, not only the internalized linguistic system, but also the mechanism in such a system is in question: what exactly are the mechanisms for L2 production (VanPatten, 2007)? Therefore, SLA researchers have diligently sought answers to which instructional efforts can promote L2 acquisition. Specifically, the SLA research will concentrate on the process or products of SLA (VanPatten, 2007). Over the years, SLA research has been developing with different underpinning theories. Next, the development of SLA research will be briefly reviewed.

The Development of SLA Research

The development of SLA is closely related to the emergence of two disciplines which appeared in the 1950s and 1960s. One is the emergence of the Audiolingual Method

supported by Behaviorism Learning Theory. The other is the emergence of children's L1 acquisition research (VanPatten, 2007). Behaviorism Learning Theory holds the view that language is composed of a series of behavior habits, and the L1 habits will interfere with L2 behavior habits. Once learners produce L2 language errors, they should correct them by imitating correct L2 language habits. Therefore, the Audiolingual Method has come into being, emphasizing reciting dialogues and practicing sentence patterns (VanPatten, 2007). However, evidence of children's L1 acquisition have shown that children can master their L1 in a very short time, which has challenged the Behaviorism Theory (VanPatten, 2007).

In the 1950s, Chomsky attacked the Behaviorism theory, claiming that SLA is a creative process, and it is impossible to acquire language through mechanical imitation and repetition (VanPatten, 2007). According to Chomsky, language acquisition is born as an innate part of human behavior and knowledge, and it can make us learn a language as quickly as children (Ding & Dai, 2008). Therefore, Chomsky regards Universal Grammar as a part of the brain. He thinks that language is not learned but developed and the human brain contains the natural elements for language knowledge (VanPatten, 2007). Due to the challenge of Behaviorism Theory, the SLA field has begun to question the teaching of Audiolingual Method, one of which is represented by Krashen's second language acquisition theory (VanPatten, 2007).

Krashen's Second Language Acquisition Theory

Inspired by Chomsky's Innate Theory, Krashen has proposed an influential theoretical model of Second Language Acquisition, which is based on five hypotheses. They are the Acquisition Learning Hypothesis, the Natural Order Hypothesis, the Monitor Hypothesis, the Input Hypothesis, and the Affective Filter Hypothesis (Liu, 2010). Among the five hypotheses, the Input Hypothesis is the core of Krashen's language acquisition theory (Tian, 2013b). In his book *Principle and Practice in Second*

Language Acquisition, Krashen (1982) has explained how learners acquire a second language in the Input Hypothesis, which also explicates how SLA happens. The Input Hypothesis can be helpful to understand SLA, which can lead to L2 production in this study.

The Input Hypothesis

Krashen (1982) emphasized that language input was the most important of Input Hypothesis. What the learners need is comprehensible input, which means that the language input should be understandable. SLA will occur when the learners can understand the input a little beyond their language level (Krashen, 1982). If the language input exceeds or falls too much below the current level of the language learners, they cannot comprehend language knowledge and SLA cannot happen (Liu, 2010). Therefore, learners can acquire language and understand the message in the target language when they are exposed to language that is slightly beyond their actual level. If the current level is "i", then the acquirer can understand input that contains "*i + 1"* (Krashen, 1982; Nunan, 2011).

As for the effective tasks in enhancing students' spoken language, teachers can design tasks that are a little beyond the students' level, which can promote the "i+1" comprehensible input of the learners. However, several studies have shown that the "comprehensible input" alone is not enough to make L2 learners achieve high levels of language acquisition (Izumi et al., 1999; Swain & Lapkin, 1995). Therefore, other perspectives in SLA will be examined.

The Output Hypothesis

To supplement the inadequacy of Input Hypothesis, Swain (1985) put forward the Output Hypothesis, which suggested that language engaging in the syntactic process of learners could foster SLA (Ellis, 2013). To explain the hypothesis, Swain (1985) investigated in a French immersion program in Canada. For students whose L1 was

English, math and science courses were all taught in the L2, French, which provided them with a great amount of comprehensible input. However, the French level of these students was not satisfactory. The students achieved high levels of French in listening and reading comprehension, but not in oral and writing skills. They had significant weakness in grammatical accuracy. Summarizing the results of this program, Swain (1985) pointed out the major reason why the students committed a lot of grammatical mistakes in French was that the language expression exercises (output) were seldom conducted. At the same time, teachers' feedback for language errors was scarce and random.

The Noticing/Triggering Function of the Output Hypothesis

Based on the above observations, Swain (1985/1995) concluded that the "comprehensible output" indeed played a crucial role in the SLA process with the noticing/triggering function.

> Noticing/triggering function: L2 learners can be aware of some language problems in their language system of production, which can trigger the consolidation of existing language knowledge or the cognitive process of acquiring new language knowledge (Swain, 1985/1995; Li, 2002).

For the noticing/triggering function, learners can pay attention to the problems in their language system through output, and then consciously analyze the language form to produce the revised output and improve the accuracy (Li, 2002). Swain and Lapkin (1995) have believed that this noticing function can trigger the output revised process, which can promote L2 learners' control and internalization of language knowledge. Therefore, with the noticing/triggering function, the language output can promote SLA.

To link with the language input, the output can enhance the effect of input through the noticing/triggering function on SLA (Swain & Lapkin, 1995), which can be illustrated in Figure 3.4.

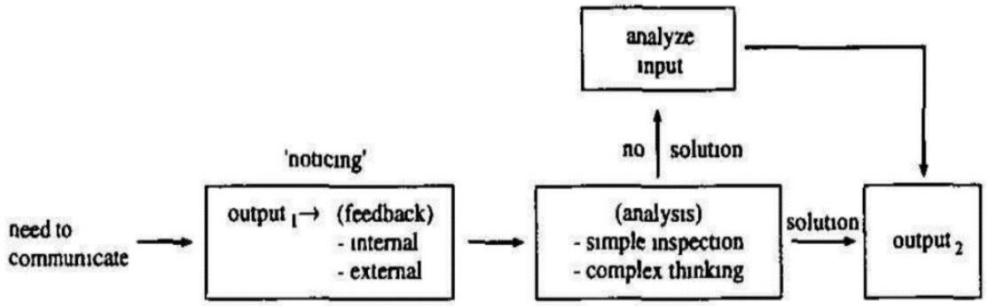


Figure 3.4: Output in Second Language Learning (Swain & Lapkin, 1995: 371)

From the output hypothesis, learners pay attention to the problems in their language system in the process of expressing meanings, which triggers the attention in the analysis of language form. The conscious attention to the language form constitutes the key link of the whole hypothesis (Li, 2002). Without conscious attention to the language form, it is impossible for learners to analyze their own language and produce the revised output for the internalization process of language knowledge and the improvement of L2 (Li, 2002). Therefore, it is indispensable for the attention and the notice function in the language output to promote SLA. During attentive noticing, the important premise for output is that the learners must have enough cognitive resources to pay attention to the form and meaning of language (Li, 2002). Thus, cognitive resources could impact on L2 output. In this study, the varied task characteristics and task conditions in the speech-making tasks can have different cognitive demands, which can influence learners' L2 production. Therefore, the cognitive perspectives influencing language output will be analyzed in the later sections. Additionally, second language production constructs will be identified first.

Second Language Production Constructs

Based on Swain's Output Hypothesis, the three perspectives of second language production have been distinguished by Skehan (1999). Complexity, accuracy, and fluency (CAF) have been viewed as the principal research variables of language

production in L2 research (Skehan, 1999). However, to review the literature of the triads, the L2 pedagogy research can be traced back to the 1980s. Brumfit (1984) was one of the earliest researchers to identify the dichotomy between fluency-oriented activities and accuracy-oriented activities. Later in the 1990s, Skehan (1989) introduced a third component of the triad, complexity, to form the CAF in the proficiency dimensions. Accuracy relates to the "degrees of deviations from a particular norm" (Housen & Kuiken, 2009: 3). Errors, compared with accuracy, is characterized as a deviation from the form (Housen & Kuiken, 2009). Fluency is about the language proficiency of a person with characterization of ease, eloquence, and smoothness of speech or writing (Housen & Kuiken, 2009). The last of the most ambiguous triad is complexity, which can be the properties of L2 performance and proficiency (L2 complexity)" (Housen & Kuiken, 2009: 4). The three constructs, complexity, accuracy, and fluency are defined in Table 3.4, which will be used as the quantitative measures in this study.

Complexity	The capacity to use more advanced language, with the possibility that such language may not be controlled effectively. This may also involve a greater willingness to take risks, and use fewer controlled language subsystems.
Accuracy	The ability to avoid error in performance, possibly reflecting higher levels of control in language.
Fluency	The capacity to use language in real time, to emphasize meanings.

Table 3.4: Definitions of Complexity, Accuracy, and Fluency (Adapted from Skehan & Foster 1999: 96–97)

The above has clarified the three constructs of L2 production: complexity, accuracy, and fluency. However, as mentioned previously for L2 speech production, there exists the cognitive factors which impact the output of L2 learners. Next, the cognitive perspectives of second language production will become the focus to explore their impact for L2 learners in this study.

3.2.3 Cognitive Perspectives of Second Language Production

To understand the cognitive factors of L2 production, how L2 speakers handle and process information in the speech production will be first explored. Then, VanPatten's (2007) Input Process Model of SLA will be depicted. At the end of this part, the cognitive factors influencing the L2 speech production will be explained along with Skehan's (2014) Limited Attentional Capacity Model to serve as the theoretical framework in this study.

The Information Process

The information process comes in three interactive stages: input, central processing, and output, which are inseparable from the basic cognition elements such as memory, attention, and integration (Skehan, 1999). For the cognitive elements, the language processing needs to take the current language signal as a clue, extract relevant language information from the long-term memory as the working memory, and integrate this extracted information into an organic part. At the same time, the extraction and integration of language processing will be regulated by attention, which is related to the working memory (Yang, 2015).

VanPatten's (2007) Input Processing Model

Implementing the influence of the working memory, VanPatten (2007) has presented the schematic Input Processing Model, which includes four steps: input, intake, developing system, and output. Among the 4 steps, there are three psychological processes: input processing, system change, and output processing (Dai & Dai, 2010). In the process of SLA, learners first get in touch with the input, then pay attention to the input, and later establish a connection between the form and meaning (Lu, 2016). After such input processing, part of the input is transformed into the intake. The intake is preserved into the working memory and may be integrated into the developing system. When the new form and meaning are absorbed, the systematic changes occur. Finally, learners will use the language materials integrated into the

developing system to produce language (Lu, 2016). The input process can be demonstrated in Figure 3.5.

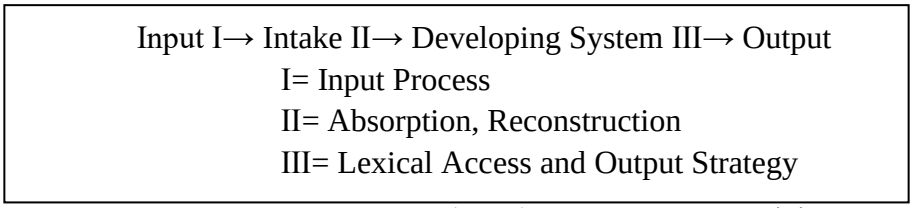


Figure 3.5: VanPatten's (2007) Input Process Model

Van Patten's (2007) model thoroughly described the L2 cognitive information process and clearly presented the dynamic process of L2 speech production. Therefore, this study will use this model as an underpinning framework to analyze the cognitive factors in task characteristics and task conditions, which will influence L2 speech production (Lu, 2016).

Cognitive Factors on Second Language Production

Considering VanPatten's (2007) model, once the connection between language form and meaning of input is established, it can be further processed and absorbed into the learner's developing system. However, not all input can be processed. Attention is the key element for learners to consciously perceive language input in order to process and internalize language knowledge (Schmidt, 2001). Therefore, in the whole process of L2 production, "attention" is a first step which is an important cognitive factor affecting oral output (Lu, 2016).

Second, in the input processing stage, the attention to language input is the result of encoding learners' working memory. Only language input encoded into the working memory can be transformed into long-term memory (Robinson, 2002). Meanwhile, absorption is first stored in working memory in the process of system changes, and then absorbed into the developing language system after integration. Therefore, when discussing the cognitive factors that affect oral production, it is necessary to

investigate working memory (Lu, 2016).

The following will discuss the effects of attention and working memory on L2 oral output separately and then describe the relationships between these two cognitive factors.

Attention

Attention is a limited, selective, and cognitive process, which is controlled by people's subjective initiative to a certain extent, and is influenced by a variety of internal and external factors, which can influence L2 production (Lu, 2016). In other words, "attention" is a necessary condition for effective input processing, which plays an important role in L2 learning (Schmidt, 2012).

In the process of L2 production, input is the first step, which is significant to the subsequent absorption, developing system, and output (Lu, 2016). Without input, there will be no output. However, not all input is of equal value. According to Schmidt's (2001) "Notice Hypothesis", only the part of input that is noticed by learners can be absorbed and processed effectively. In order to learn the linguistic features in the target language, learners must consciously pay attention to the linguistic features in the input. During L2 language learning, learners need to pay attention to how certain language form and meaning in the target language, which is paying attention not only to the external features of the language form, but also to the meaning of the language and situational uses. Only by paying constant attention to the relationship between form and meaning, can we learn to use language in communication (Lu, 2016). Such attention to language form and meaning can trigger absorption and contribution to the development of learners' interlanguage, which is the language system produced by learners. With the learners' interlanguage developed, their language output will be reached. As far as the oral output is concerned, the function of attention is to guide learners to prioritize meaning with sacrifice and less

resources to language form (VanPatten, 1990).

Therefore, attention itself is a restricted psychological process, which is influenced by a variety of internal and external factors. Skehan (1998) believes that the cognitive mechanism of attention is influenced by frequency, salience, instruction, individual differences in processing ability, learners' interlanguage status, and task demand.

In general, "attention" plays a coordinating role between the input and the memory system. The noticed information is more likely to enter the memory system and lay the foundation for the change of interlanguage (Lu, 2016), promoting the development of L2 production. However, when attention is limited, L2 speech production is influenced.

The Working Memory

As mentioned above, attention plays an important role in the initial stage from input to output. However, as a limited selective cognitive process, "attention" is limited by the capacity of working memory (Sawyer & Ranta, 2001). The working memory has two functions: temporary information storage and information processing. Daneman and Carpenter (1980) believed that the storage and processing abilities of individuals in cognitive activities were limited by the working memory capacity. In fact, through the whole process of oral production, the working memory capacity affects not only "attention", but also the whole process of input processing, system change, and output. Learners with large working memory capacity enjoy the advantage in the cognitive processes. When processing the input language materials, they can allocate the cognitive resources more effectively and take fuller account of the different aspects of form and meaning (Lu, 2016). In the process of language production, the cognitive resources can be redistributed to other stages of language output, such as the conception of speech contents (Wen, 2007).

In short, working memory capacity has a very important constraint on L2 production, which influences the stages of input processing, system change and CAF of oral output. However, working memory capacity is a relatively stable cognitive mechanism for learners, which cannot be changed, arbitrarily. Therefore, teachers can reduce the workload of working memory and the restrictive effects of working memory for L2 production (Lu, 2016). In this study, the workload of the working memory in task characteristics and task conditions can be reduced to view their effects on learners' L2 production.

Attention and Working Memory

As mentioned above, attention and working memory play influential roles in the different stages of information process in L2 production. However, both cognitive factors do not exist in isolation, but interact with each other. Their interaction determines the outcome of L2 production (Lu, 2016). Between them, "attention" plays a key role in the whole language production process, which determines the information to enter the memory system and to be absorbed, thus triggering the change of the interlanguage system, and improving the output. Meanwhile, "attention" is limited by working memory capacity, and learners with large working memory capacity are more likely to consider the different aspects of language form and meaning to produce better language (Lu, 2016). In short, L2 production is influenced along with the interactive, cognitive factors of attention and working memory (Lu, 2016).

3.2.4 The Limited Attentional Capacity Model: Trade-off Hypothesis

Based on the limited resources of attention and working memory, it is natural for the learners to focus on meaning instead of form in a communicative context (VanPatten, 2007). The possibility could be that form is lost at the expense of advanced language, to achieve the primary goals of fluency, and meaningful expressions (Skehan, 2014).

Focusing on the cognitive perspectives, Skehan (2014) has emphasized the values of attention and working memory and proposed the assumed framework which indicates that learners' cognitive capacity of attentional working memory resources is limited to achieve speech production in CAF. This assumption will be adopted as the theoretical framework in this study.

To further explain the Limited Attentional Capacity Model, Skehan (2014) has argued that L2 learners cannot focus on all the three language production constructs because the ability to process information is limited and has demanded more attention to resources. Consequently, a better performance in one of the constructs which is obtained with the sacrifice of another (Li & Fu, 2018). Therefore, it is very likely that there are trade-off effects in CAF as language learners struggle to conceptualize, formulate, and articulate the oral message in Levelt's model of speech production (Ellis, 2013). Among CAF, there is the initial distinction between focus on form and focus on meaning/fluency (Skehan & Foster, 2001). According to VanPatten's (1990) proposal, focus on form and meaning will come into significant competition with one another due to the lack of the capacity of cognitive resources from L2 learners (Xu & Chen, 2017). Therefore, when the capacity of L2 learners' cognitive attention is committed to meaning or fluency, the performance of a task would come with effective and real-time communication at the cost of attention to form with less complex and accurate language (Skehan & Foster, 2001). However, when the limited capacity of L2 learners' cognitive resources is concentrated on form, they can produce more accurate language on one hand, and more complex, challenging, and difficult language on the other hand (Xu & Chen, 2017), but with less fluent language. The tension between the two dimensions can be seen in Figure 3.6.

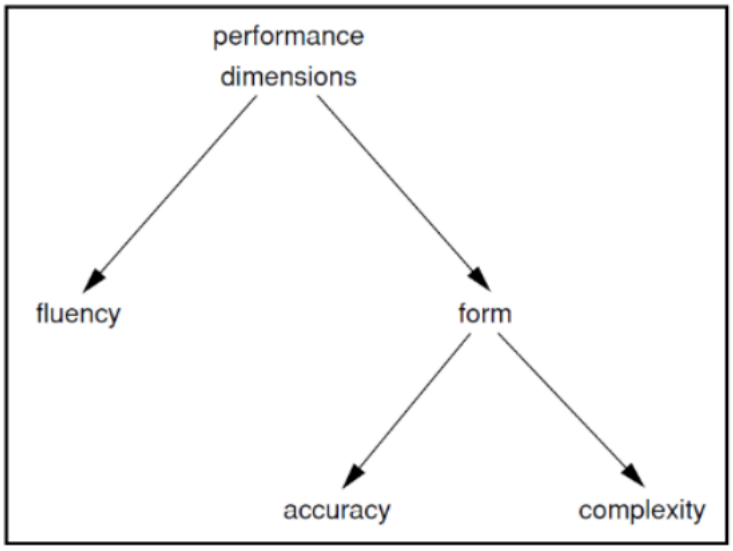


Figure 3.6: Theorizing Dimensions of Task Performance (Skehan & Foster, 2001: 190)

Skehan and Foster (2001) have further manifested the trade-off effects. Complexity comes along with the reconstruction of language learning process, concerning "the learner's willingness to use more challenging and difficult language" (Skehan & Foster, 2001: 190). When learners recognize that "their interlanguage systems are limited and require modification, they are more likely to use more complex language, but also to attempt to pressure their own language systems" (Skehan & Foster, 2001: 191). That is to say, the greater complexity means "a willingness to experiment, and try to extend or make more elaboration on the underlying interlanguage system" by having less accuracy in the language (Skehan & Foster, 2001: 191). On the contrary, the greater accuracy "reflects a greater degree of conservatism, as the learner tries to achieve greater control over more stable interlanguage elements" with less complex language (Skehan & Foster, 2001: 191). The trade-off effects from Skehan's Limited Attentional Capacity Model can be summarized in Table 3.5.

This study has introduced spoken language production with Levelt's (1989) model of L1 speech production. After the clarification of L1 speech production, the L2 production with the elaboration of SLA has been explored. Then, the triad

measurements of speech production have been identified as CAF. After that, the cognitive perspectives of the information process are demonstrated. Then, the cognitive factors of attention and working memory are explained based on VanPatten's (2007) Input Processing model. In the end, Skehan's (2014) Limited Attentional Capacity Model is proposed as the theoretical framework for this study.

Skehan's Limited Capacity Model: trade-off effects in language production	Focus on meaning/fluency: F ↑; Focus on form: (A ↓, C↓) Focus on form: (A ↑, C↑); Focus on meaning/fluency: F ↓
	Focus on form: (C ↑, A ↓); (A ↑, C ↓)

Table 3.5: Skehan's (2014) Limited Capacity Model: Trade-off Effects in Language Production

The above are the reviews of two separate themes: tasks and spoken language production. Next, the relationship between tasks and spoken language production will be explored, specifically, focusing on task characteristics, task conditions, and their influences on learners' L2 speech production.

3.3 TASKS AND SPOKEN LANGUAGE PRODUCTION

For the third theme, task characteristics, task conditions, and their influence on L2 speech production will be reviewed. After that, the research gap will be identified to justify this research. Finally, the conceptual framework will be demonstrated.

3.3.1 Task Characteristics and Spoken Language Production

In the task design, Skehan (1999) has identified the relationship between task characteristics and spoken language performance. The task features include familiarity of information (concrete-abstract; familiar-unfamiliar material) and degree of structure (structured-unstructured). For the familiarity of information, the tasks "vary as to whether they require information which is familiar to the participants as

part of their personal experience" (Skehan, 2011: 235). For the degree of structure, some tasks appear with a clear and over-arching structure while some do not (Skehan, 2011). When assessing language production in terms of familiarity of information, Skehan (2011) has proposed that tasks with concrete or familiar information will "advantage accuracy and fluency" (83). Less attention and working memory are needed for the speakers to assemble the speech and more attention is given to ensure accuracy and fluency. For degree of structure, it is argued that accuracy will also be impacted (Skehan, 2011). The tasks with clear and sequential structure will promote the speech production in freeing the attentional working memory resources and direct towards the goals in producing accuracy (Skehan, 2011). The task characteristics and their relationships with spoken language production can be summarized in Table 3.6.

Task Characteristics	Familiarity of Information	Degree of Structure
Spoken Language Production	Familiarity↑ A ↑ F ↑ C-	Structure↑ A ↑

Table 3.6: The Relationship between Task Characteristics and Spoken Language Production (Skehan, 2011)

Familiarity of Information

It was proposed by Skehan (2011) that the familiarity of information could promote fluency and accuracy of task performance. The familiar knowledge of information can ensure less demands on attention and working memory to access information quickly and easily for the conceptualization of speech. No significant change was found in complexity as the speakers paid attention to the complex language to express familiar events (Skehan, 2011).

Empirical research has been done to test this hypothesis. Foster and Skehan (1996) have compared the effects of language performance on the personal information exchange, narrative, and decision-making tasks. It is indicated that these three tasks demand has disparate levels of attentional resources along with gradually less known

and expected information, which would require the progressing workload of cognitive attention and working memory to impact on language production (Foster & Skehan, 1996). For familiarity of information, the personal task can be compared with the other two. Foster and Skehan (1996) have indicated that the personal task achieves greater accuracy than the narrative task. The same goes with fluency, as the personal task generates fewer pauses and higher fluency than the other tasks (Foster & Skehan, 1996).

Degree of Structure

Skehan and Foster (1997) have suggested that tasks with structure are linked with more accuracy. Such influence was made from the findings of Levelt's (1989) model of speech production that tasks with a certain structure would initiate higher accuracy in language production. Drawing upon Levelt's (1989) model of L1 speech production, Skehan (2009) has suggested that the foundation of more accurate speech is based on the formulation stage. The formulation stage involves the lemma retrieval and the morpho-syntax building processes. This stage is correlated with speech errors, pausing and hesitation phenomena, and slips of the tongue, which are the elements of speech accuracy and fluency (Skehan, 2009). Therefore, the more structured tasks enable learners to possess clearer macrostructure of speech formulation, thus allocating more attentional resources and working memory capacity in speech production (Skehan, 2009).

Nevertheless, the task characteristics of familiarity of information and degree of structure alone do not simply influence language production. How the tasks are manipulated in language classroom can also impact on task and L2 production (Ellis, 2013). Next, the task implementation conditions will be examined.

3.3.2 Task Conditions and Spoken Language Production

During task implementation, the task planning conditions are related to L2 production.

For the spoken language, it will include the planning process during which the speakers can determine the contents of the speech with the linguistic devices selected (Ellis, 2005). The spoken language planning could contain various levels of discourse plans, sentence plans, and constituent plans to be blended in the articulation of the speech (Clark & Clark, 1977). Therefore, the task conditions of planning could be influential to L2 production. For the task conditions, Ellis (2005) has generated the principal types: pre-task planning and within-task planning. The types of planning can be seen in Figure 3.7.

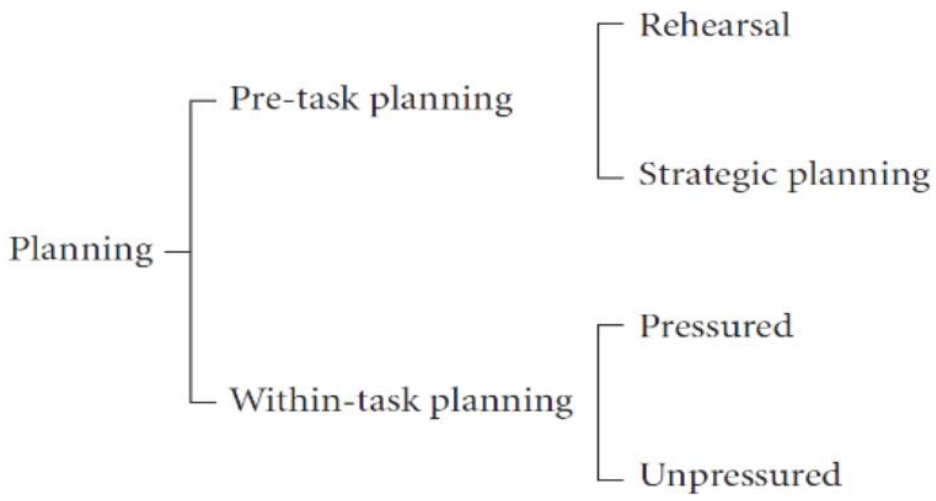


Figure 3.7: The Principal Types of Task-based Planning (Ellis, 2005)

The distinctive features between pre-task planning and within-task planning are that pre-task planning can plan the task before task performance, while within-task planning proceeds online immediately when the task is performed (Ellis, 2005). The pre-task planning includes rehearsal and strategic planning. Rehearsal enables the learners the chance to prepare and perform the whole task before task performance, while strategic planning entitles speakers the opportunities to consider the contents of language production (Ellis, 2005) (see Table 3.7). What differentiates the strategic planning from other types of pre-task planning involves brainstorming the speech contents, examining the model performance of a task, and searching the dictionary. As

for within-task planning, there are pressured (tasks performed within a limited amount of time) and unpressured (tasks performed without the limited amount of time) tasks (Ellis, 2005). For the speech-making tasks, the focus will be on pre-task planning which can allow more time for the students to prepare and ensure more pedagogical interventions from the teachers than within-task planning. Next, the pre-task planning and the influence on L2 production will be displayed.

Pre-task planning is planning that is done before learners perform a task.	Rehearsal	Planning takes the form of an opportunity to perform the complete task once before performing it a second time.
	Strategic planning	Planning includes the content to be expressed and the language to be used but without an opportunity to rehearse the complete task.

Table 3.7: Pre-task Planning and Types of Pre-task Planning (Ellis, 2009: 474)

Pre-task Planning and Spoken Language Production

The abovementioned Levelt's (1989) model of speaking involved three processes: conceptualization, formulation, and articulation, which is in line with the view that a learner possesses limited cognitive resources and has the difficulty in handling the whole aspect of language production at the same time (Ellis, 2009). This can provide the basis for the task planning studies which shows that the types of pre-task planning can alleviate the pressure of a learner's attentional resources and working memory. Task planning affects the competition and the trade-off effects in CAF of L2 production (Ellis, 2009). Based on previous findings, Ellis (2013) has demonstrated the effects of pre-task planning for language production in Table 3.8. For this research study, the strategic planning and rehearsal will be adopted to examine their influence on L2 production in CAF.

Language Performance	Rehearsal	Strategic Planning
Fluency	No effect	Positive effect
Accuracy	No effect	Effects sometimes evident
Complexity	No effect	Positive effect

Table 3.8: The Effects of Rehearsal and Strategic Planning on Language Performance (Ellis, 2013: 133)

Strategic Planning and Spoken Language Production

As stated above, strategic planning contributes to conceptualization based on what has to be communicated and has a positive effect on fluency and complexity but does not have a significant effect on accuracy (Ellis, 2013). Next, previous studies of strategic planning on complexity, accuracy, and fluency will be reviewed.

For complexity, clear results of strategic planning have been proposed that planners can have more complex language production than non-planners. Cookes (1989) explained that more complex sentences and the wider range of lexical words were displayed in the 10 minutes strategic planning during task performance. Forster and Skehan (1996) echoed the findings that detailed planners generated more considerable subordination than the non-planners. Ortega (1999) demonstrated that it was higher for the complexity measure, the mean number of words per utterance, under planning conditions. Wang and Song (2015) reported an empirical study on the influence of strategic planning on L2 oral performance, which aimed to explore whether different length of preparation time (0, 1, 2, 3 minutes) led to differences in L2 oral performance, and whether there was an interaction between preparation time and language proficiency. The results showed that learners' syntactic complexity increased.

For accuracy, the influences of strategic planning are varied. Some studies have revealed the enhancement of accuracy for strategic planning, while some have implied no effect. Ellis (1987) described that more accurate grammatical use of past tense was found in strategic planning. Meanwhile, the significant increase of accuracy was indicated from the 1-minute planners than the non-planners (Mehner, 1998). However, Yuan and Ellis (2003) specified that no significant effects were found in strategic planning for accuracy. Nevertheless, other studies suggested that promotion of accuracy from strategic planning could be found in some tasks or in some conditions (Ellis, 2005). For the different types of tasks, it was demonstrated that greater

accuracy from strategic planning was found on personal and narrative tasks aside from the decision-making tasks (Skehan & Foster, 1997). For some conditions, Foster and Skehan (1999) found that accuracy of strategic planning was greater when it was teacher-led. In the previous research mentioned above, Wang and Song (2015) found that accuracy remained stable, when the preparation time reached 3 minutes.

For fluency, studies have shown that strategic planning has beneficial results. Foster and Skehan (1996) suggested that the speakers have less silence and stop less compared to no planning in the 3 tasks they proposed. Ortega (1999) supported the view with the evidence that speakers can speak faster for the speech rate for L2 Spanish in the story-telling task when planning strategically. Meanwhile, Yuan and Ellis (2003) described the positive effects of fluency in strategic planning. Li and Fu (2018) investigated the comparative effects of strategic and unpressured within-task planning on L2 Chinese oral production and the role of working memory in mediating the effects of the two types of planning. The results revealed that strategic planning enhanced fluency.

Rehearsal and Spoken Language Production

From Ellis's (2013) analysis, rehearsal has shown no significant effects on CAF of language performance. It is suggested that rehearsal is beneficial to the successive language performance of the same task. However, the beneficial effect cannot transfer to a different task even though it is the same type as the original task (Ellis, 2005). Next, previous studies of rehearsal in CAF will be examined.

By comparing one learner's retelling of a Tom and Jerry cartoon, on two separate occasions, three days apart, Bygate (1996) indicated that rehearsal promoted complexity, with the support of more lexical verbs, more regular past tense forms, a wider range of vocabulary, and cohesive devices from the learner.

Another unique task that involved rehearsal was devised by Lynch and Maclean (2000), in an ESP course, to prepare the medical profession for English presentation. A “poster carousel” task was designed for students to read an academic article and prepare a poster presentation based on it. Students then stood in front of their posters for viewing and questions from other groups. Each stand would receive 6 visitors for the same questions. Therefore, substantial opportunities were given for rehearsal. The study found that the recycling output could enhance both greater accuracy and fluency for the learners (Lynch & Maclean, 2000).

An additional study was reported by Bygate (2001) to investigate the effects of practicing specific types of tasks on both a second performance of the same task and on performance of a new task of the same type. Greater fluency and complexity were manifested on the second performance. Therefore, it is helpful to practice that particular type of task. Nevertheless, the practice cannot help the performance of a new task of the same type, which means there is no transferrable effect of rehearsal to the same types of a new task.

Zhong (2021) explored the impact of task complexity and pre-task planning types on oral output. The results have shown that rehearsal can improve students’ fluency, but it is not helpful to complexity and accuracy.

The above has explored the task characteristics, familiarity of information (concrete-abstract; familiar-unfamiliar material) and degree of structure (structured-unstructured) and task conditions of strategic planning and rehearsal in pre-task planning and their influences on L2 production. For this study, familiarity of information, degree of structure, strategic planning and rehearsal will be implemented to find out their actual effects for spoken language production of CAF in English Listening and Speaking in a university of science and technology in China. Moving towards this study, the research gap will be justified.

3.4 RESEARCH GAP

Looking at the Chinese context in recent years, researchers have been working on task characteristics (familiarity of information; degree of structure) and task conditions (pre-task planning of strategic planning and rehearsal) for L2 learners. Previous studies have suggested two variables with their influence on L2 production (Bui, 2014; Luo, 2009; Skehan & Shum, 2014; Wang, 2014; Tan & Dong, 2007; Xu, 2015). However, research on these four variables together is limited in L2 oral production in the designated research context.

Tan and Dong (2007) explored topic familiarity and pre-task planning on the accuracy of speech production for English majors in China. By investigating the topic familiarity through 2 types of tasks: task 1 (retelling an unknown story) and task 2 (impromptu speech: telling an embarrassing personal experience) and the variation of pre-task planning (no preparation for task 1, 3 minutes preparation for task 2), it was indicated that task 2 with more familiar topic and preparation time would have higher language output in accuracy.

Luo (2009) analyzed the effects of various teaching activities on the topic familiarity and strategic planning for middle school students in China, which reduced students' cognitive and language pressure. In such cases, their language production was improved.

Bui (2014) researched on the effects of topic familiarity, strategic planning for university students in Hong Kong, China. The results showed that both topic familiarity and strategic planning promoted more fluent language, but strategic planning was a stronger form as indicated by its effect sizes.

Skehan and Shum (2014) investigated video-based narrative retellings with the variables of degree of structure and the online processing conditions for 2nd year

university students in China. It was suggested that the more structure and less pressured online processing conditions could generate more accuracy and complexity.

Wang (2014) studied strategic planning and rehearsal and gathered data from 77 undergraduates (L1: Chinese and L2: English) in Hong Kong, China, by shooting a video narrative task in English. It was signified that strategic planning promoted speech complexity and fluency, which showed the conceptualization stage in speech production enhanced language complexity and fluency. In addition, speech complexity, fluency, and accuracy were developed by rehearsal, with the support of improved speaking quality.

Xu (2015) examined the effects of strategic planning and structure on language output of the oral tasks. A quantitative quasi experimental study was conducted on the oral output for English majors in China. The results showed that: 1) when completing nonstructural tasks, the preparation time ensured complexity in language output; 2) under the condition of strategic planning, oral fluency was promoted.

In conclusion, no significant research can be found for exploring the four variables, topic familiarity, degree of structure, strategic planning, and rehearsal, together for L2 speech production in China, let alone in the context of English Listening and Speaking course in a university of science and technology. Therefore, this study will fill the research gap.

3.5 CONCEPTUAL FRAMEWORK

To draw a conceptual map (see Figure 3.8) of what have been discussed, so far, the context of the study targets at ELT in China. Then, the research angle is focused on CE in China, specifically, the course of English Listening and Speaking, in which the speech-making tasks are chosen.

To underpin the research question: how do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking, this Literature Review chapter investigates the three themes: "Tasks", "Spoken Language Production", and "Tasks and Spoken Language Production".

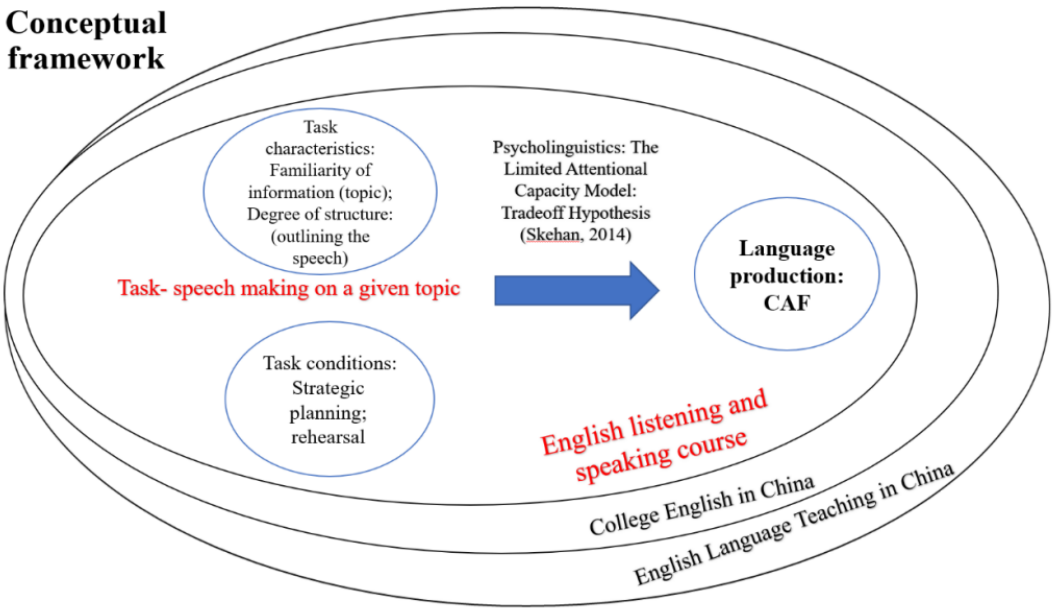


Figure 3.8: Conceptual Framework of this Study

For the first theme "Tasks", the conceptions, definitions, task types and the speech-making tasks were reviewed. Then, the second theme "Spoken Language Production" was investigated. For L1 speech production, Levelt's (1989) model of first language speech production was demonstrated. Researching the L2 context, the SLA field was reviewed along with spoken language production constructs: CAF. After that, the cognitive perspectives of information process were demonstrated along with the cognitive factors of attention and working memory. Based on VanPatten's (2007) Input Process Model, Skehan's (2014) Limited Attentional Capacity Model was proposed as the theoretical framework of this study. Then, in the last theme "Task and Spoken Language Production", the relationships between tasks and spoken language production were elaborated. For tasks characteristics and spoken language production, familiarity of information and degree of structure were examined for their influences

on spoken language production. For task conditions, strategic planning and rehearsal in pre-task planning were chosen for their impacts on spoken language production.

Then, the research gap was identified to justify this study. In the end, the conceptual framework was shown for the study of tasks and spoken language production. Next, Chapter 4 will further explore the methodology and design of this study.

Chapter 4 Methodology

In this study, the research question is “How do different task characteristics and task conditions impact students’ spoken language production (CAF) in English Listening and Speaking?”. This methodology chapter will be divided into three sections to seek answers to the research question. First, there will be a discussion of philosophical paradigms with the selection the pragmaticism paradigm. Then, the application of Mixed Methods Research (MMR, including definition, purpose, and the selection of research design; and a justification of the explanatory model of the MMR design in this study) will be explained. Finally, a research design along with the details of research plan, data collection arrangements, data analysis methods and analytic framework of this study will be devised.

4.1 RESEARCH METHODOLOGY

As the current study belongs to the field of social sciences, the paradigms and paradigm wars in social sciences will be discussed. To resolve these wars, pragmaticism could be the solution. The pragmatist values what works in the research study. Therefore, the task characteristics, task conditions, and how they facilitate students’ spoken language will be investigated.

Broadly speaking, research methodology is often understood in terms of specific methods when conducting research. The hypotheses behind research study are also considered as vital components in research methodology (Paris & Reynolds, 1983). These basic logical patterns or hypotheses are referred as the paradigms that determine the starting point, essence and method of a research, and the philosophical foundation behind them (Guba, 1990). To understand the research methodology of this study, the comprehension of paradigms in social science research will be explored.

4.1.1 Conceptualization of Paradigm-Where It Begins

The term "paradigm" is originally popularized by Thomas Kuhn in his landmark book, *The Structure of Scientific Revolutions* (1962). In this book, paradigm is defined in two ways, one as the beliefs of values of a certain group of researchers, the other as the plan to solve research problems (Kuhn, 1962).

> On one hand, it stands for the entire constellation of beliefs, values, and techniques shared by the members of a given community. On the other, it denotes one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal science.

Sharing similar views as Kuhn, scholars have identified and discussed paradigms, in recent decades. Paradigms are a series of basic beliefs guiding people's daily life, behaviors, and discipline research, or ways to guide the researchers' outlook of the world and reveal the complexity of our realities (Guba, 1990). Probing further into the term "paradigms", Morgan (2007: 49) has described them as the "systems of beliefs and practices that influence how researchers select both the questions they study and methods they use to study them".

As for the nature of paradigm, Guba and Lincoln (1994) have described it as a basic belief system which contains a trilogy of ontological, epistemological, and methodological assumptions. The ontological question explains the nature of reality, describing what can be known and how things really are (Guba & Lincoln, 1994). The epistemological question studies "what the nature of the relationship between the knower or would-be knower is and what can be known". The methodological question explores "how the inquirer (would-be knower) finds out whatever he or she believes can be known" (Guba & Lincoln, 1994: 108). These three fundamental questions can be regarded as the major foci in the analysis of paradigms (Guba & Lincoln, 1994). Therefore, the research paradigm chosen in this research study will be mainly

distinguished from the three layers of philosophical paradigm (ontology, epistemology, and methodology).

As this study of task characteristics and task conditions on students' spoken language production belongs to the social science research, the paradigms in social science research will be explicated for further selection in this study.

4.1.2 Paradigms in Social Science Research

Traditionally, there are two controversial parties of paradigms in social science research. One is the objectivist approach of positivism/empiricism, and the other one is the subjectivist approach of constructivism, or hermeneutics/phenomenology (Deng & Pan, 2002).

The rise of the first camp is positivism in the 19th century, which has a profound impact on the social sciences including sociology, psychology, and pedagogy. In this tradition, people hope to follow the scientific model to establish social sciences (Zhu, 2002). The social sciences have three basic influences which stem from positivism. First, methodology purism indicated that all mature knowledge fields are not disparate in essence. Meanwhile, the natural and social sciences are not different in methods (Zhu, 2002). It advocates the application of the research methods which have achieved great success in natural sciences to social sciences. Many positivists believe that only by applying the methods of natural sciences to social sciences can social sciences really become science and people can really acquire "knowledge" in this field (Zhu, 2002). Second, positivism links knowledge with experience and takes a phenomenological and empiricist viewpoint. Only the phenomenon that can be observed and studied can become knowledge (Zhu, 2002). Third, the aim of social sciences is to seek the universal law of society. Therefore, positivism always occupies an important position in the debate about nature, purpose, and scientific status of

social sciences (Zhu, 2002). In the debate regarding the scientific nature of social sciences, “quantification” and “objectivity” are the two key words (Zhu, 2002).

Summarizing the above 3 principal foci of philosophical paradigm from Guba and Lincoln (1994), the ontology of objectivist positivism approach belongs to realism, believing that there are “objective facts” dominated by the laws of nature. Therefore, objectivism belongs to positivism in epistemology and contains the belief that it is not subject to the values, scenarios, and time of the researchers (Deng & Pan, 2002). The methodological choices of positivism are nomothetic, deductive, and based on quantity, which tends to reveal the causal relationship between things and determine whether or how much one variable causes the change of another, to obtain an explanation and prediction that can be extended (Burrell & Morgan, 1979). The role of the researcher in the positivist study is to observe and measure in a neutral way, avoid intervening or influencing the research object, and pursue the “objectivity” of research to the maximum extent. The positivism paradigm and quantitative research methods have been the main body of social science research from the 19th century to the 1960s (Deng & Pan, 2002).

Turing to the 1960s, there has been growing interests in qualitative research in the field of social sciences. Qualitative research has risen with the development of anthropology and ethnography and their use in social science research (Zhu, 2002). The 1950’s to 1970’s was the golden age for the development of qualitative research. In this period, scholars began to make qualitative methods standardized and rigorous. It was also at this time that the legitimacy of qualitative research was widely recognized (Zhu, 2002). Hermeneutics has gradually replaced the empirical method and become the main paradigm to reveal social phenomena and the human experience (Deng & Pan, 2002). To explain further, the hermeneutics paradigms refer to the reconstruction of different subjective meanings that people construct in a particular situation and how these meanings relate to each other to form a whole interpretation

and understanding (Greene, 1990). Those who hold the hermeneutic paradigm believed that there were various subjective facts in the human psychological structure.

Moving to the three key foci of philosophical paradigm, the subjectivist approach is rooted in nominalism ontology. Its epistemological stance lies in anti-positivism, which assumes that the subject and object of cognition are inseparable, and the research findings are the results of the interaction between them (Burrell & Morgan, 1979). The methodological decision is idiographic and the corresponding qualitative research method is inductive in nature (Burrell & Morgan, 1979). It starts from concrete observations and gradually constructs general models and concepts where the observation about what is happening in a certain situation is rooted in direct experience, not through prior assumptions or deductive reasoning (Deng & Pan, 2002). "Subjectivity" has replaced "objectivity" as an important feature of qualitative research. Therefore, researchers' values and perceptions play an important role (Patton, 1990).

4.1.3 Paradigm Wars in Social Science Research

How It Began

Facing two contentious parties of paradigms, the wars have begun with a series of debates, mainly the nature and relative epistemological value of positivism and anti-positivism (Zhu, 2002).

On one hand, there are various objections to the study of positivism research in social sciences. For example, supporters for the anti-positivism think that quantitative research cannot deal with ideology, thoughts, and feelings; they are accused of treating people as objects, ignoring their personal characteristics, and forcing them to fall into different categories (Zhu, 2002). Hermeneutics holds that there are two fundamental errors in the premise of large-scale quantitative research: first,

ontologically speaking, it assumes that people will be completely driven by the causal structure. Second, from the perspective of methodology, human behavior cannot be expressed by only one variable, because any aspect of human behavior is affected by a series of uncertain localized meanings (Zhu, 2002). Apart from the above arguments, there is still a lot of resistance to the anti-positivism qualitative research from the camps of positivism. Researchers who adopt a qualitative approach are labeled as the "soft scientists" from their accusers. Their work is said to be unscientific or just exploratory, or completely personal and full of prejudice (Zhu, 2002).

The disputes of positivism/quantitative and anti-positivism/qualitative research in the field of social sciences are labeled as the "paradigm wars", which stem from the distinctive perceptions of incommensurability and incomparability of paradigms for quantitative and qualitative research (Bryman, 2006). In the 1970s and 1980s, the "paradigm wars" were, fiercely, on, which has shown serious antagonism and hostility between the two camps (Zhang, 2011). Denzin (2010: 421) has echoed the view of Bryman (2006) that the disputes of the 1980s paradigm wars lie mainly in the principal understanding of the nature of the two paradigms in incompatibility and incommensurability:

> (a) Quantitative and qualitative methods were fundamentally different; (b) interpretive or theoretical paradigms could not be combined; (c) there is no value or theory-free inquiry (d) paradigms are incommensurable and (e) methods have incompatible assumptions, meaning they cannot be combined.

In the 1980s, the paradigm wars were concluded with the demise of positivist quantitative research and the flourish of interpretivist, qualitative research in social science research (Gage, 1989).

In conclusion, in the paradigm wars, both paradigms have their shortcomings in social science study. The positivist quantitative research conducted by means of empirical investigation and experiment has been criticized because it does not consider that the

people studied in social sciences are different from the materials studied in natural sciences (Chang & Luo, 2005). Interpretivist, qualitative research, which relies on participatory observation and in-depth investigation, has also been criticized for its lack of scientific credibility. In fact, both methods have their own shortcomings, but also have their own advantages (Chang & Luo, 2005).

For this research study, a compromise of methodological approaches with the combination of strength and weakness of both paradigms will be chosen for the philosophical paradigm. This methodological design is viewed as pragmaticism, which can signal a resolution of the paradigm wars.

Pragmaticism-the Possible Ending of the Paradigm Wars

In the paper "The Paradigm Wars and Their Aftermath", Gage (1989) has expressed a possible way to end the battles. "The answer to the future lies with us…Pragmatic, philosophical analysis shows us the foolishness of these paradigm wars and the way to an honest and productive rapprochement between paradigms" (Gage, 1989: 10). Therefore, the pragmaticism paradigm could put an end to the wars and be the reconciliation for the two parties.

One of the important reasons for the possible cessation of paradigm wars is the criticism of the distinctive incompatibility and incommensurable nature of the two paradigms by scholars in the field of social science research, that is, the simple opposition between positivism quantitative research and interpretative qualitative research is wrong (Zhang, 2011).

To suppress the opposition, Bryman (2008) has attempted to make some practical speculations on the combination of both paradigms into research investigations. First, positivism takes the position of empiricism. The same goes with the participatory observation, which belongs to anti-positivism qualitative research. Participatory

observation, like positivism quantitative research, is also empirical research when it establishes a connection between the observed categories; Second, qualitative research sometimes allows researchers to participate in some causal processes which are very similar to quantitative research and which seeks causal statements (Bryman, 2008); Third, positivism requires sociological research to go through the process of putting forward a hypothesis theory and then testing the hypothesis like natural sciences, so that people will think that quantitative research is acceptable for testing the hypothesis theory, while qualitative research is not. However, qualitative research, such as participatory observation, is appropriate for testing hypothesis theory (Bryman, 2008). From the above considerations, both the positivist research and the interpretivist research are not strictly in opposition. They could be in convergence in social science research and work for the advancement of research studies.

Therefore, for this research study of task characteristics and task conditions and their influence on students' spoken language production, the combination and convergence of both paradigms will be adopted as it could avoid the opposition of both paradigms and bring out the strengths of both. The qualitative research can be used to study subjective opinions and experiences, while quantitative research can obtain quantitative data. Both methods can create knowledge and increase understanding of the world (Chang & Luo, 2005). Thus, the current study will combine both methods for research inquiry. To scrutinize the combination of the two paradigms, the pragmatism paradigm will be explored.

4.1.4 Pragmatism Paradigm

Pragmatism as a Philosophy

Dating back to the understandings and works of Peirce, William James, and Dewey, pragmatism centers on the practical application of a concept in scientific investigation (Robson, 2011). To define truth as "what works" in its effect, pragmatism, is

considered as the alternative paradigm, and is endorsed with eclecticism and pluralism, questioning the dichotomy between positivism and interpretivism and seeking a combination of quantitative and qualitative methods. It can be argued that both paradigms are not differentiated at the epistemological or ontological level and that many commonalities are shared in their research approaches to inquiry (Feilzer, 2010).

However, pragmatism is also one of the most misunderstood lines of thought in modern, western philosophy. When it comes to pragmatism, most people will think of it as being eager for quick success and instant benefit, which is both opportunistic and unscrupulous (Zhou & Li, 2009). Therefore, pragmatism philosophy is regarded as a utilitarian philosophy, a "philistine philosophy" or "vulgar philosophy". The main cause of such misreading lies in people's understanding of truth (Zhou & Li, 2009).

To argue what truth is, Peirce, the founder of pragmatism philosophy, has put forward the truth theory of utility principle. James has developed the truth view as "utility is truth". Dewey has further enhanced the truth viewpoint of instrumentalism based on James' view (Zhou & Li, 2009). In Dewey's understandings, ideas, and theories are all tools that people use to make their actions successful. Therefore, whether ideas and theories can make people's actions successful or not becomes the criterion of whether they are truth (Zhou & Li, 2009). Truth is a tool. Like all the other tools, its value lies not in itself, but in the effect that it can produce (Zou, 1990). As tools, truth only matters when it is applicable, easy to use, or not (Zhou & Li, 2009). Thus, the major understanding of truth for pragmatism refers to what works best in effect in research inquires regardless of seeking a lot of success or constant success in the study. So, pragmatism should be considered as a research philosophy in scientific research instead of the utilitarian philosophy by its accusers.

The Nature of Pragmatism

The nature of philosophical paradigm is described by Guba and Lincoln (1994) as a

trilogy of ontology, epistemology, and methodology. Added to the tripartite concepts of Guba and Lincoln (1994), Biddle and Schafft (2015) have summed up four fundamental concepts of philosophy of knowledge. 1) Ontology is the nature of existence, reality, and the understanding of it (Hammersley, 2012). 2) Epistemology refers to "the nature of the relationship between the knower and would-be knower and what can be known" (Guba & Lincoln, 1994: 108). 3) Axiology is the ethics, values, and beliefs we hold. 4) Methodology is the rationale of the methods we select to investigate our research (Biddle & Schafft, 2015).

To explain pragmatism as a philosophical paradigm, it is described as "recognizing the existence and importance of the natural and physical world as well as the emergent social and psychological world" in ontology. Epistemologically, knowledge is regarded as "being constructed and based on the reality of the world we experience and live in". Axiologically, pragmatists adopt "an explicitly value-oriented approach to research from cultural values, which specially endorses shared values such as equality and progress" (Robson, 2011: 28-29). Methodologically, mixed methods, with pragmatism as the philosophical partner, consolidate world views, research questions, methods, inferences, and conclusions (Creswell & Tashakkori, 2007).

Pragmatism, as the foundations of mixed methods research, is asserted by Morgan (2007), Creswell and Plano-Clark (2011), as a philosophical worldview. Multiple worldviews might be applied in mixed methods research to the research design. For practical reasons, the very worldview which works best for the research will be chosen. The worldview model by Creswell and Plano-Clark (2011), indicates that the selected design typology is tailored to specific inquires (Birddle & Schafft, 2015). It is suggested, in the essence, that the worldview model is as follows:

> The normal science of those using mixed methodologies is a pragmatic approach to their research, selecting either an explicitly pragmatic justification for their work or pragmatically selecting a worldview based on what difference it makes to do so (Birddle & Schafft, 2015: 326).

To conclude, pragmatism pushes aside the disputes between quantitative and qualitative advocators and puts an end to the paradigm wars, which indicates that what matters the most is "whether the research can find out what the researcher wants to know" (Hanson, 2008: 109). It serves as "normal science" for mixed methods research, whose values embody an acknowledgement of both quantitative and qualitative research methods and "offer a chance to produce a properly integrated methodology for the social sciences" (Feilzer, 2010: 14).

For this study, the fitness for teaching and learning purpose matters the most. Philosophically, I would consider myself as a pragmatist. Therefore, to explore the answer for the research question "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", pragmatism will be adopted as the worldview and the philosophical paradigm. Mixed methods research will be chosen as the methodology in this study. Next, further elaboration of mixed method research will be revealed.

4.2 RESEARCH METHOD-MIXED METHODS RESEARCH (MMR)

As this study is based on mixed methods research from the pragmaticism paradigm, this section will first explain the definition of MMR. Then, the research purposes of MMR will be justified. Finally, the MMR research design will be selected.

4.2.1 Definition of MMR

By examining 19 leaders in the fields of MMR, Johnson et al (2007: 118-123) generalize 5 themes from the definitions of MMR (see Table 4.1). Theme 1 refers to the contents of the quantitative and qualitative research. Theme 2 indicates the time and space for the quantitative and qualitative data collection and analysis. Theme 3 symbolizes the breadth of the broad continuum of worldviews. Theme 4 explains the

reasons and purposes of the MMR study. Theme 5 points out the approach and direction of the study (Johnson et al, 2007).

Theme 1	Contents	Quantitative + Qualitative; Qualitative + Qualitative; Quantitative + Quantitative
Theme 2	Time and Space	Data collection stage; Data collection and data analysis stages; All the stages of research
Theme 3	Breadth	A broad continuum from one's definitions to one's worldviews
Theme 4	Reasons and Purposes	1) Provide a fuller picture and deeper understanding; 2) Validate and explicate findings from another approach and gain confidence in conclusions 3) Provide richer/more meaningful/more useful answers to research questions.
Theme 5	Orientations	Bottom-up approach; Top-down approach; Bottom-up/Top-down conceptualization continuum

Table 4.1: Five Themes of Mixed Methods Research (Johnson et al, 2007)

With the above-mentioned generalization of the five themes, the definition of MMR can be specified. Formally defined as the research where quantitative and qualitative research are intellectually and practically synthesized, Mixed Methods Research is regarded as the "third wave", "the third methodological choice/moment/movement" or "the third paradigm for research" (Johnson & Onwuegbuzie, 2004; Denscombe, 2008). Denzin (2010: 422) illustrates an interesting metaphor which compares MMR to "Wag the Dog". Because several variations on a common theme are generated from the mixed methods movements, the methods are like the dog's tail, wagging from one variation to the other.

4.2.2 Purposes of Mixed Methods Research

As for the purposes of MMR, Greene, Caracelli, and Graham (1989) have generated five specific purposes in social science studies: triangulation, complementarity, development, initiation, and expansion. The purposes can be linked to the current

study respectively.

1) Triangulation aims for convergence and collaboration of quantitative and qualitative results, which drives different methods to study the same phenomenon. In this study, the quantitative method and qualitative method will be combined to analysis students' spoken language production for both quantitative and qualitative results in the conclusion.

2) Complementarity seeks elaboration, enrichment, and enhancement of the study, resulting from different methods (Greene, Caracelli & Graham, 1989). The quantitative and qualitative data and findings will be complementary, elaborated, and enhanced in this study.

3) Development means one method (quan) can help inform the other method (qual). The quantitative and qualitative data can be cross-applied to confirm the results of this study and form conclusions.

4) Initiation is to discover paradoxes, contradictions, and new perspectives for the improvement of studies (Greene, Caracelli & Graham, 1989). The quantitative and qualitative result may contradict each other in this study, which could initiate the need for further discussion.

5) Expansion echoes Thomas Cook's multiplism to expand the scope and breadth of research with different methods for different components (Johnson et al, 2007). In this study, the scope and breath will be developed and enhanced as both methods generate different data for data analysis.

Each of the above purposes will be beneficial in yielding answers of the research question "How do different task characteristics and task conditions impact students'

spoken language production (CAF) in English Listening and Speaking?", which justify the rightful use of MMR. Next, the discussion of mixed methods design will be explained.

4.2.3 Mixed Methods Design

How to Choose MMR Design

In line with Creswell and Plano-Clark (2007), it is suggested that there are three major decisions to make before selecting a particular type of mixed methods design. Three major reasons are outlined in the decision tree of selecting a particular research design (Figure 4.1).

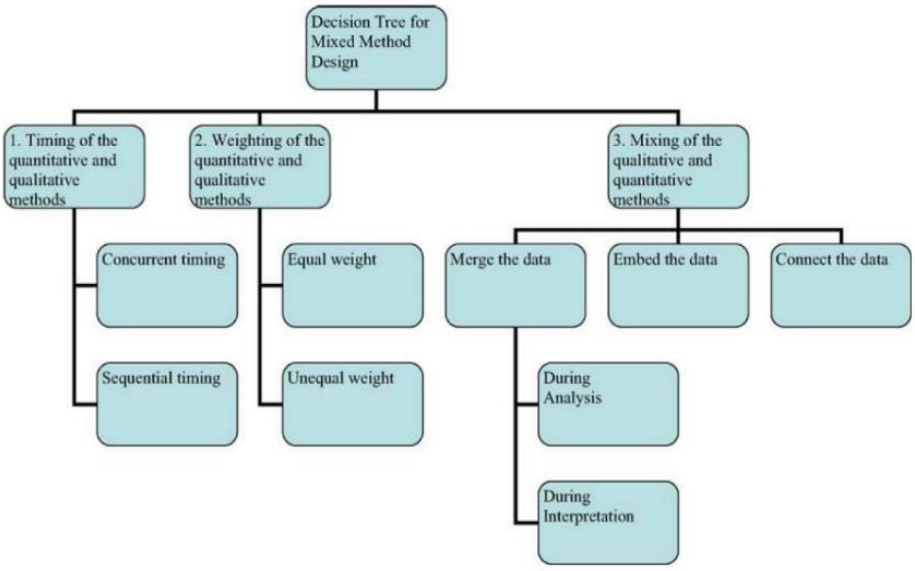


Figure 4.1: Decision Tree for Mixed Methods Design (Creswell et al., 2003; Creswell & Plano-Clark, 2007)

First is the timing of quantitative and qualitative designs, which differs from concurrent timing and sequential timing. The second reason is the weighting, or the importance of quantitative and qualitative method, considering whether they are of equal or unequal weight. The third is where the mixing of the qualitative and

quantitative methods will happen (Creswell, et al., 2003; Doyle et al., 2016). Bearing in mind the three major reasons for selecting mixed-method design, researchers can select from the four major types of MMR as follows (Creswell & Plano-Clark, 2007) (Table 4.2).

Design Type	Variants	Timing	Weighting	Mixing	Notation
Triangulation	Convergence Data transformation Validating quantitative data Multilevel	Concurrent: Quan and Qual at the same time	Usually, equal	Merge the data during the interpretation or analysis	QUAN+QUAL
Embedded	Embedded experimental Embedded correlational	Concurrent or sequential	Unequal	Embedded one type of data within a larger design using the other type of data	QUAN or QUAL
Explanatory	Follow-up explanations Participant selection	Sequential: Quantitative followed by qualitative	Usually, quantitative	Connect the data between the two phases	QUAN→QUAL
Exploratory	Instrument development Taxonomy development	Sequential: Qualitative followed by quantitative	Usually, qualitative	Connect the data between the two phases	QUAL→QUAN

Table 4.2: The Major Mixed Methods Research Types (Creswell & Plano-Clark, 2007: 85)

What to Choose for MMR Design?

Having summarized 12 major classifications of mixed designs from 1989 to 2003, Creswell and Plano-Clark (2007) conceptualize 4 major types of mixed methods, the Triangulation Design, the Embedded Design, the Explanatory Design, and the Exploratory Design. The Triangulation design is a one-phase design where qualitative and quantitative methods are carried out at the same time with equal importance. It

helps to collect different but complementary data for better understanding of the research question (Creswell & Plano-Clark, 2007). The Embedded Design embeds one set of methods within the other set of methods, in the design level. It means one method, for instance, quantitative methods, is the major design, while qualitative is in a supportive and subsidiary role. The Explanatory Design plans, a two-phase mixed methods design, using qualitative data to "help explain or build upon the initial quantitative results" (Creswell & Plano-Clark, 2007: 71). The Exploratory Design, based on the necessity of an exploration when there are no available measures, instruments, or guiding framework, is best used to explore a phenomenon, which begins with the qualitative method, and then continues with the quantitative method, to yield quantitative results with the interpretation of qualitative method (Creswell & Plano-Clark, 2007) (Table 4.2).

For the design of MMR in this study, the timing of the quantitative and qualitative methods will be sequential. The weight of importance of the quantitative and qualitative designs will be unequal with the quantitative analysis in the major weight and qualitative explanations of participants in the following. Then, there will be connection of data analysis of the two methods. Therefore, the explanatory model of triangulation will be suitable to conduct this research based on the above three decisions. In this study, the conceptual framework for task characteristics and task conditions in spoken language production is based on the CAF measures. So, the quantitative data of CAF measures will be collected for the speech tasks. As for students' perceptions of varying features in task characteristics and task conditions and their influence on language production, qualitative data will be collected through questionnaires and focus group interviews. The quantitative data will be connected before the qualitative data. Meanwhile, both data will be merged to analyze and interpret the research findings. Therefore, the explanatory model of triangulation will be chosen (Figure 4.2).

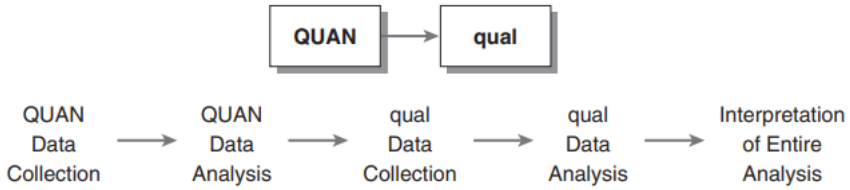


Figure 4.2: The Explanatory Model of Triangulation (Creswell et al., 2003)

The explanatory model selected in this study can provide comprehensive perspectives in viewing the research issues. The quantitative data can produce numeric findings for the influence of task characteristics and task conditions on spoken language production. In addition, the qualitative data will probe for deep insights from students' viewpoints of the research question (Creswell, 2014).

The triangulation of data can provide multiple perspectives of data as well as increasing the validity of data (Creswell, 2014). However, the explanatory model also poses potential challenges for data analysis, as both quantitative and qualitative data would consume time to analyze. Problems might occur when the quantitative and qualitative findings differentiate in the data analysis process (Creswell, 2003). But thorough research design and strict research procedures will be presented and arranged to ensure the completion of this research.

4.3 RESEARCH DESIGN

Based on the mixed methods design, this section will clarify the research design along with the details of the research variables, research plan, data collection arrangements, quantitative data, and analysis, qualitative data, and analysis. In the end, the analytic framework will be summarized in this study.

4.3.1 Research Variables

To explore the answer to the research question "How do different task characteristics

and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", a research design has been proposed to explore the two themes: first, task characteristics: familiarity of information, degree of structure. Second, task conditions: strategic planning and rehearsal for students' L2 speech production (Table 4.3).

Task: speech making on a given topic			
Research phases	**Phase 1**	**Phase 2**	**Phase 3**
Task characteristics	Concrete: a familiar topic	Abstract: an unfamiliar topic	Concrete: a familiar topic
	Unstructured: without the speech outline	Unstructured: without the speech outline	Structured: with the speech outline
Task conditions	Strategic planning	Strategic planning & Rehearsal	Strategic planning & Rehearsal

Table 4.3: Research Design

This research focuses on alteration of the variables in task characteristics and task conditions. After the data collection, their influences on spoken language production will be found from both quantitative and qualitative data.

First, the task type selected is speech-making. As mentioned in Chapter 3, the speech-making tasks are chosen to be the outlining and the impromptu speech genres to ensure purposeful, authentic communication, and open-ended outcomes. In the classroom practice, meaningful and real-life communication from the learners is practiced through brainstorming for ideas and negotiating meanings. When the speeches are done in class, there is meaningful communication among students to exchange viewpoints on the speech topics. Second, the task characteristics contain

two variables: familiarity of information and degree of structure. Familiarity of information means whether the information is concrete or abstract for speech topics, that is, whether the students are familiar with the speech topics. The degree of structure has been varied with the structured and unstructured tasks. The structured task provides the practice of writing speech outline before speaking. The unstructured task does not provide this practice. Based on Skehan's (2011) views, tasks with concrete or familiar information can facilitate accuracy and fluency. The structured tasks can increase accuracy in language production (Skehan, 2011). This study investigates the varied familiarity of information and degree of structure in the specific research context to see their influence on spoken language production.

For the pre-task planning, strategic planning and rehearsal have been selected to conduct the research. Strategic planning includes the following: 1) Students can brainstorm the speech contents; 2) The teacher provides examples of model speeches and glossaries. Rehearsal means students can practice the whole speech, once, before speech-making. Considering Ellis's (2013) hypothesis, rehearsal has no effect on CAF, but it is suggested that rehearsal is beneficial on the successive language performance of the same task (Ellis, 2005). Meanwhile, strategic planning has shown positive effects on fluency and complexity, but does not have a significant effect on accuracy (Ellis, 2013). In this study, answers have been sought to those assumptions. In this research study, three phases have been designed. In each phase, the task characteristics variables and pre-task planning conditions have been changed to reveal their influence on students' spoken language production. The quantitative and qualitative data have been analyzed and discussed with the previous assumptions.

4.3.2 Research Plan

Overall Sampling Participants and Strategies

Three classes of students (freshmen in XXX university) from the English Listening

and Speaking course were chosen as the sampling of the research participants because they were accessible as the researcher's students at that time in the language lab for the speech recordings. This course lasted for 16 weeks from Sept. 21st, 2020 to Jan. 7th, 2021. Each class had approximately 40 students. Altogether, there were about 120 students from seven different majors, participating in this study. One convenience sample group of 30 students was selected with gender differences of both male and female from different majors and levels of English proficiency, as the group of participants were available for quantitative speech recordings. The 120 students from the 3 classes were accessed for the qualitative questionnaires. As for the focus group interviews, the sample distribution characteristics such as different majors and gender were taken in consideration. Among the six participants in the focus group, there were 50% of male and female students respectively with the average age of 18, from the major of business management, computer science and environmental science and engineering. Therefore, this study adopted the cluster sampling method of classes. The interviewees were two students from each of the three classes. In the first focus group interview, all the six students participated. In the second one, five students were interviewed.

The three phases are shown in Table 4.3. For each phase, students were asked to make a speech on a given topic, and all speeches in the three phrases were recorded for the quantitative data. Phase 1 began from Week 7 of the course. Every student made a speech based on a familiar topic without the structured outline, along with strategic planning. The speech topic for phase 1 can be seen in Table 4.4. Phase 2 was conducted, one week later, on Week 8. The phase 2 task was abstract without the speech outline, along with strategic planning and rehearsal. After Phase 2, the online questionnaire and focus group interview were conducted. Then, Phase 3 was one week after phase 2. The phase 3 task was familiar, structured with strategic planning, and rehearsal. After Phase 3, the online questionnaire and focus group interview data were collected again.

Topic familiarity	Phase 1 Concrete topic: How do shared bikes influence our life? Please speak for 1 min.
	Phase 2 Abstract topic: Describe a trip that you have taken in Shenzhen. Please speak for 1 min.
	Phase 3 Concrete topic: In recent years, social networking such as QQ and WeChat, in China, has become very popular. Do you think social networking pushes people closer or further apart? Use specific reasons and/or examples to support your views. Please speak for 2 min.
Degree of structure:	Phase 1 and 2, unstructured Phase 3 structured 1. Introduction 2. Body 3. Conclusion
Strategic planning	Brainstorm for the speech contents; provide examples of model speeches; provide glossaries
Rehearsal	Students will be able to perform the speech once before the task.

Table 4.4: Research Contents

4.3.3 Ethical Issues

When doing the mixed-method research, ethical approval was obtained from the University of Nottingham Ningbo China and informed consent was obtained from the students-participants in the English Listening and Speaking course (Appendix 1). The nature of the research was explained to the participants so that they could decide whether they would participate in the study (Plowright, 2011). They were granted the right to withdraw from the research at any time, without penalty. Confidentiality and anonymity were ensured so that students were protected from revealing their names in the speech recordings, questionnaires, and focus group interviews in data analysis (Cohen, Manion & Morrison, 2011).

4.3.4 Data Collection Arrangements

Phase 1: Week 7

1. One concrete topic without a speech outline was selected, along with strategic

planning.

2. 30 students were selected with gender differences, from the 3 classes.
3. Quantitative data collection: CAF measures were applied to the sample group's recordings.

Phase 2: Week 8

1. The teacher provided an abstract topic, unstructured task with strategic planning, and rehearsal in Phase 2.
2. Quantitative data collection: CAF measures were found from the sample group's recordings.
3. The first online questionnaire was collected for the students' perceptions between phase 1 and 2 for their opinions on task characteristics and task conditions among the 3 classes for about 120 students.
4. Focus group interview: the first focus group interview was conducted among the sample group to compare the results between phase 1 and 2 for students' perceptions on the research issues.

Phase 3: Week 9

1. Phase 3 was conducted with a concrete topic, structured, strategic planning, and rehearsal.
2. Quantitative data collection: CAF measures were found from the sample group's recordings.
3. The second online questionnaire was done to collect the students' perceptions among phase 1, 2, and 3 for their opinions on task characteristics and task conditions.
4. Focus group interview: the second focus group interview was conducted among the sample group to compare the results between phase 1, 2, and 3 for students' perceptions on the research issues.

4.3.5 Challenges for Data Collection

The challenges for data collection can be found in 3 aspects: pilot study, quantitative data collection and qualitative data collection.

Pilot Study

In the research design section, familiarity of information for task characteristics was not easy to confirm. As the researcher, I could not determine what kinds of topics were familiar to the students. Therefore, a pilot study was conducted a few weeks earlier, from the data collection, to test the students' views on whether the topics were familiar.

The same problem occurred for task conditions. This research was supposed to study strategic planning, alone, for task condition. However, in the pilot study, it was found that students cannot produce a speech without any pre-task planning. Therefore, rehearsal was added along with strategic planning, as task conditions, to vary their influence on spoken language production in this study.

Quantitative Data Collection

During the pilot study, I was unable to collect all the students' recordings, as I could not control the class time adequately. So, among the 3 classes, only 2 classes of their recordings were collected. Then, in the next few weeks, I adjusted the teaching plan and moved the data collection at the beginning of the lesson to ensure complete collection of the recordings.

Qualitative Data Collection

For the first focus group interview, I collected the opinions from all the 6 students. However, in the second interview, one student did not show up. Therefore, I had to settle for 5 students in data analysis.

4.3.6 Quantitative and Qualitative Data and Analysis

For this mixed methods research, both quantitative and qualitative data have been analyzed after data collection. In this section, the details of both data and the analytic approach will be clarified.

Quantitative Data and Analysis

For quantitative data, the sample group's recordings were collected and analyzed among the 3 phases. Next, the sample group and their recordings will be displayed.

Sample Group Recordings

The sample group contained 30 first-year students with majors ranging from automation, architecture, computer science, business management, civil engineering, and environmental science and engineering. Among the 30 students, there were 16 male and 14 female students, with each 10 of them from 3 English classes. In the three samples of 10 students, their oral English scores varied from advanced, upper-intermediate, and intermediate based on the marks of final exam in the oral speech test of this course. The 30 students along with the 3 English classes were recorded in 3 research phases for varied task characteristics and task conditions. In Phase 1, 30 recordings were collected for a concrete and familiar topic with their views on the topic of "How do shared bikes influence our life?". The speeches were unstructured with strategic planning such as brainstorming for the speech contents; providing examples of model speeches and glossaries. One week later in Phase 2, there was an abstract topic of "describing a trip in Shenzhen" without the speech structure. After brainstorming the ideas in strategic planning, 30 recordings were collected with unstructured speech topic. Then, after rehearsal, 30 more recordings were collected in Phase 2. In Phase 3 one week later, there were 30 recordings for a familiar topic on students' opinions of "social networking and interpersonal relationships" with the speech structure and strategic planning. After rehearsal, 30 more recordings were collected with structured, and the same familiar topic in Phase 3. Therefore, there

were 150 MP3 recordings of 30 students in the sample group of 3 research phases for quantitative data analysis.

The quantitative data in this study are the CAF measures of spoken language production from the sample group's 150 recordings in all 3 phases. Among the 3 phrases, the task characteristics (Concrete/Abstract; Unstructured/Structured) and task conditions (Strategic planning/Rehearsal) have been varied to identify their influence on learners' spoken language production on CAF. The quantitative data of the CAF measures have been analyzed in SPSS to answer the research question "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking". Next, the oral data analysis will be examined.

Analyzing Oral Data

Complexity, Accuracy, and Fluency (CAF)

To analyze the oral data for spoken language production, the definitions of CAF will be revisited. Complexity refers to what extent the learners can elaborate on the language. The extent comes in two aspects: one is the learners' willingness of using the challenging and difficult language; the other is the readiness of using various structures in spoken language production (Ellis & Barkhuizen, 2005). Accuracy means "how well the target language is produced in relation to the rule system of the target language" (Skehan, 1996: 23). Fluency is depicted as "the capacity to produce speech at a normal rate, without interruption, or as "the production of language in real time without undue pausing or hesitation" (Ellis & Barkhuizen, 2005: 139). The CAF measures will be adopted in the quantitative analysis in this study. For the quantitative analysis, the analytic units of the oral data for this study should be chosen for the analysis of CAF measures.

Units for the Measurement of Spoken Language

As for the analysis of oral data, Foster et al. (2000) have suggested that researchers need to seek a unit, or a segmentation of oral data in quantitative analysis. It is argued by Foster et al. (2000) that the units can be semantic or intonational. However, the semantic and intonational units are found to be unreliable. For semantic units, it is not easy or possible to analyze the exact meanings of an idea or argument with certainty (Foster et al., 2000). As for intonational units, the pausing and intonational features for L2 learners are viewed as questionable since those features can be altered for non-native speakers. To search for more reliable and valid measurement for oral data, the syntactic units are easier to identify in the oral data and are regarded as more desirable for the oral data analysis (Foster et al., 2000). So, in this study, the syntactic units will be analyzed.

The general and popular syntactic units are introduced as T-unit, C-unit, and AS-unit (Foster et al., 2000). T-unit is revealed as the most popular unit for both written and spoken language analysis (Foster et al., 2000). To get to know T-unit, the full name is minimal terminable unit. It means a main clause and all clauses that belong to the sentence (Wang, 1985). Nevertheless, inadequacy was found for the T-unit analysis as some of the recorded speeches were filled with hesitation and repetition. The dysfluent language and incomplete sentences cannot be analyzed with the T-unit (Foster et al., 2000). For such problems, Young (1995) has made some modifications for data analysis by adding substantial footnotes of speeches for the insufficient information of oral data. These modifications are later called the communication unit, C-unit, which includes the T-unit along with the elliptical nature of spoken language (Foster et al., 2000). However, from Bygate's (1988) analysis, the T-unit and C-unit analysis were found unsatisfactory for the speech of independent noun phrase satellite units, which are common for L2 learners for their L1 as the topic-comment language. In this study, the sample group students (L1, Chinese, topic comment language; L2, English), will make speeches regarding the abstract and familiar topics. Therefore, we should search for another valid and reliable unit for the oral data analysis. For dealing

with the oral data, Foster et al. (2000) have put forward the Analysis of Speech Unit (AS-unit). The AS-unit is specified as:

> …a single speaker's utterance consisting of an independent clause or sub-clausal unit, together with any subordinate clause (s) associated with it (Foster et al., 2000: 365).

On one hand, the independent sub-clausal units are very common in oral speech. Differing from T-unit, the AS-unit includes the independent sub-clausal units which are phrases that can be put into a full clause with the completion of the elliptic elements. The AS-unit can elaborate the nature of independent sub-clausal units more clearly of the oral data than C-unit analysis (Foster et al., 2000). On the other hand, the subordinate clause contains at least one finite or infinite verb along with one other clausal element (Subject, Object, Complement). Since subordination is a crucial element for the complexity measure, the AS-unit can be useful in measuring syntactic complexity (Foster et al., 2000).

An example of the AS units from Phase 1 Student A's recording is as follows. The number of AS-unit is counted as 5 (Appendix 3).

Student A Speech Recording

Without fluency features:

> 1. As a as a new transportation, shared bikes changed our lives greatly. 2. But a new way for journey, it is convenient and cheap, and so on. 3. But it problems exist, too, exist, but problems exist as either. 4. The illegal parking The when that slims, the safet who the safet who who save the car? 5. And the the QR code were the QR code were were were plagiitces, influence , change , our life. Yeah.

After explaining the units of oral data analysis, the investigation on the selection of CAF measures will be reviewed, respectively, for this research study.

CAF Measurements

A wide range of specific measures have been adopted by researchers for the quantitative data analysis. It is suggested that the decisions regarding the CAF measures should be research driven or data driven (Ellis, 2013). Next, the detail measurements of the CAF construct for this research will be presented.

Complexity Measures

Complexity refers to the extent of language produced by learners (Ellis & Barkhuizen, 2005). Skehan (2003) defines the language complexity to be the extent of the produced language within learner's interlanguage system. Further developing the concepts of complexity, Bulté and Housen (2012) believe that language complexity is mainly composed of lexical and grammatical complexity. The advantage of this dichotomy lies in the formation of specific dimensions of language complexity from the two basic elements of language: vocabulary and grammar, which makes language complexity visible and palpable, and enables researchers to establish measurements of these two dimensions (Liu & Miu, 2018).

Lexical Complexity

Lexical complexity, specified to the lexical features, has variety in lexical sophistication including rareness, diversity, and density (Yang, 2014). According to Zheng and Feng (2017: 60), the three features of lexical complexity can be described in the following ways:

1) Lexical sophistication is the proportion of low-frequency words in speaking. For analysis, the online software VocabProfile (website: www.lextutor.ca) was adopted. The software divides text words into four lists: the most commonly used 1000 words, the less commonly used 1001-2000 words, academic vocabulary, and vocabulary out of the three lists. In this study, the academic vocabulary and the vocabulary out of three lists were used to calculate the Beyond 2000 value, which represented the complexity of vocabulary.

2) Lexical diversity, which is the ratio between parts of vocabulary (type) and token (TTR), and the Uber index, was used in this study. Lexical diversity was calculated by the following formula: (Square of logarithm of the token)/ (logarithm of the token-logarithm of the type).

3) Lexical density, the proportion of the content of words in each speech to the total number of words (Content Word Ratio, CWR), was calculated online by VocabProfile.

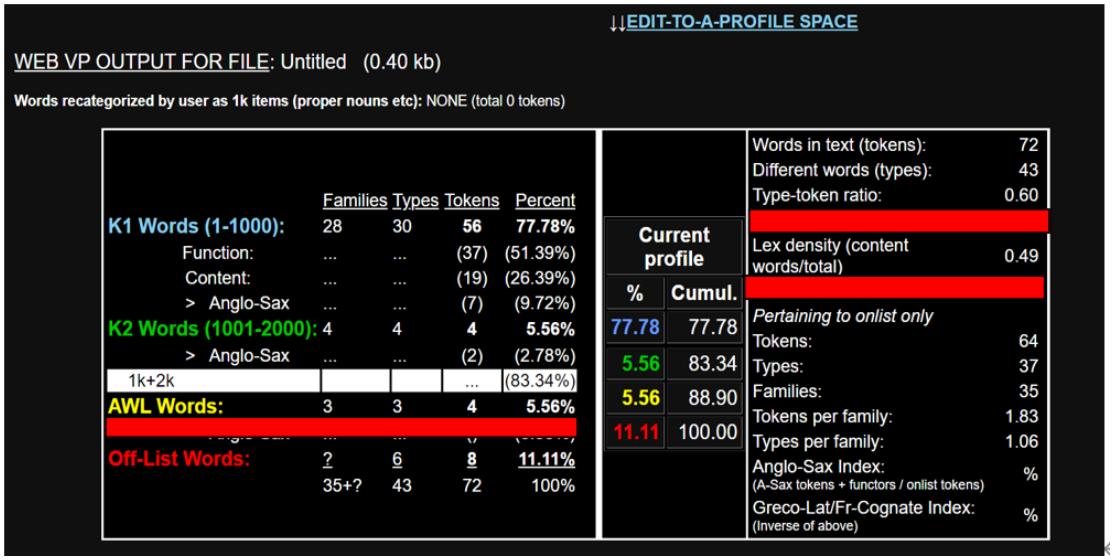
An example of lexical complexity can be seen in Appendix 3:

P1 Student A's Speech Recording

Without fluency features:

> As a as a new transportation, shared bikes changed our lives greatly. But a new way for journey, it is convenient and cheap, and so on. But it problems exist, too, exist, but problems exist as either. The illegal parking The when that slims, the safet who the safet who who save the car? And the the QR code were the QR code were were were plagiitces, influence , change , our life. Yeah.

P1 Student A's lexical complexity is marked in the following table:



↓↓EDIT-TO-A-PROFILE SPACE

WEB VP OUTPUT FOR FILE: Untitled (0.40 kb)

Words recategorized by user as 1k items (proper nouns etc): NONE (total 0 tokens)

	Families	Types	Tokens	Percent
K1 Words (1-1000):	28	30	56	77.78%
Function:	...	...	(37)	(51.39%)
Content:	...	...	(19)	(26.39%)
> Anglo-Sax	...	...	(7)	(9.72%)
K2 Words (1001-2000):	4	4	4	5.56%
> Anglo-Sax	...	...	(2)	(2.78%)
1k+2k			...	(83.34%)
AWL Words:	3	3	4	5.56%
Off-List Words:	?	6	8	11.11%
	35+?	43	72	100%

Current profile	
%	Cumul.
77.78	77.78
5.56	83.34
5.56	88.90
11.11	100.00

Words in text (tokens):	72
Different words (types):	43
Type-token ratio:	0.60
Lex density (content words/total)	0.49
Pertaining to onlist only	
Tokens:	64
Types:	37
Families:	35
Tokens per family:	1.83
Types per family:	1.06
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

Lexical Complexity

AWL Words: 5.56%

TTR: 0.60

Lexical density: 0.49

Grammatical Complexity

For grammatical complexity, Bulté and Housen (2012) believes that it consists of grammatical richness and depth, which can be divided into syntactic richness and depth, morphological richness, and depth. Syntactic richness and depth refer to the variety and change of phrases, clauses, sentences, and other syntactic forms (Liu & Miu, 2018). Morphological richness and depth refer to the variety and difficulty of words in their derivations (Liu & Miu, 2018). In view of Bulté and Housen (2012), syntactic complexity is a vital component for grammatical complexity, but whether morphological complexity can be measured remains questionable as it is difficult to judge the applicable conditions of rules and meaning caused by affixes or prefixes without corresponding measurements, at present. So, in this research, the syntactic complexity of students' oral production was measured.

Syntactic Complexity

The syntactic complexity refers to the "complexity of constructions used in sentences" (Yang, 2014: 45), which can be measured by the clauses per unit (Xing, 2014). Since the speech tasks in this study are all the oral data of students' recorded speeches, the AS-unit will be selected as explained above.

For syntactic complexity, it contains both the overall complexity index and the specific complexity index. Generally speaking, T-unit average length, C-unit average length, and AS-unit average length are the overall indicators of syntactic complexity (Liu & Miu, 2018). However, these overall indicators have some limitations, which can be easily confused as the syntactic structures with different features and functions. Take the average length of T-unit as an example, the basic idea is that the longer the structure is, the more complex it is (Liu & Miu, 2018). However, it could ignore the characteristics and functions of some specific syntactic sentences (Biber et al., 2016).

Because the length of a sentence does not necessarily mean complexity, the long structure can be composed of relatively simple words and syntactic elements, while the short structure can also contain difficult collocations and words (Liu & Miu, 2018). In terms of syntactic features and functions, some specific indices have been proposed, which mainly measure the complexity of sentences, through clauses and phrases (Liu & Miu, 2018). At the clausal level, specific measurement indicators include clause/AS-unit, clause/C-unit, clause/T-unit, subordinate clauses/clauses (Liu & Miu, 2018).

In terms of measurement for syntactic complexity, the existing measurement indicators have their own advantages and disadvantages. Not only do the overall indicators need to be combined with specific indicators, the researchers should also choose the indicators suitable for their own research according to research purposes, participants, and scope (Liu & Miu, 2018).

Therefore, based on the initial analysis of the pilot data, the speech sentences in this study tend to be broken and short. So, to ensure the validity and reliability of this study, the syntactic complexity was measured by the combination of the general indicator, the average length of AS-unit and specific indicators clause per AS-unit, to find out the different task characteristics and task conditions and their influences on spoken language production.

The same example of syntactic complexity for Phase 1 Student A can be found as follows:

Syntactic Complexity

The average length of AS unit: Total words/ AS-unit 74/5=14.8

Clause per AS unit: Clauses/AS-unit 13/5=2.6

Accuracy Measures

Polio (2001) has described the linguistic accuracy as "a broad term that generally has to do with the absence of errors" (94). In this study, errors are counted as pronunciation errors as well as grammatical errors. Wolfe-Quitero et al. (1998) have given a detailed examination of the measurements of accuracy. They argue that the best indicator of accuracy is the percentage of error-free clauses, which is measured by calculating the number of learners' errors in speech. The error-free clauses ratio is the ratio of the number of correct clauses to the total number of clauses. For the unfocused tasks, Ellis and Barkhuizen (2005) have recommended the use of overall indicators, such as the ratio of error-free clauses, or the number of errors per hundred words. As for our study, the speech tasks are all unfocused. What we strive to find out is the learners' ability to produce correct speech with varied task characteristics and task conditions. However, the error-free clauses index alone is viewed to be problematic as elision (the omission of a sound or syllable when speaking) is common in the speech data. This problem can be solved with the additional measure of errors per hundred words (Ellis & Barkhuizen, 2005). Therefore, the students' spoken language production has been analyzed by the ratio of error-free clauses and the number of errors per hundred words for accuracy.

The same example of accuracy for Phase 1 Student A can be found as follows:

Accuracy:

The ratio of error-free clauses: the numbers of error-free clauses/total numbers of independent clauses, subclausal units, subordinate clauses *100

2/15*100=13.33

The number of errors per hundred words: Errors/total words/100

6/74/100=0.00081

Fluency Measures

Fluency is usually related to native-like spoken language (Xing, 2014). It is divided into two major types. One type is the temporal variable, including the speed of

speaking. The other is the hesitation phenomena of dysfluency. The major temporal variables are speech rate and number of pauses, while dysfluency elements of the hesitation phenomena involve false starts, repetitions, reformations, and replacements (Ellis & Barkhuizen, 2005). After doing a pilot analysis of students' recordings, both the temporal variables and the hesitation phenomena can be spotted distinctively. Therefore, both indices will be considered and analyzed in this study. Temporal variables involve mainly the speech rate and number of pauses for students' fluency in speech-making (Ellis & Barkhuizen, 2005). The speech rate is measured by the number of words per minute on the speech task, where the number of words (leaving the dysfluencies) is calculated and divided by the total number of minutes of the speech (Ellis & Barkhuizen, 2005). The number of pauses is counted as the total number of filled and unfilled pauses of the learners. As for hesitation phenomena, they can include false starts, repetitions, reformations, and replacements (Ellis & Barkhuizen, 2005). False starts mean incomplete utterances in the speeches. Repetitions refer to words and phrases that are repeated without any modifications, while reformations mean repeated words and phrases, with some modifications. Replacements in the oral data show that the lexical items in the speeches are replaced by other items (Skehan & Foster, 1999). As false starts are not distinctive in the speech recordings, they are not counted in this study.

The same example of fluency for Phase 1 Student A can be found as follows:

Fluency:

Speech rate: 74 words/minute

Pauses and fillers: 12

Repetitions: 8

Reformations: 0

Replacements: 1

Statistical Analysis

Regarding the statistical analysis, the CAF measures of spoken language production have been studied with different task characteristics and task conditions for SPSS analysis in this research. Based on the data from the recordings, the change of CAF production will be influenced by the varied task characteristics and task conditions.

On one hand, by viewing the research design of Phase 2 and Phase 3, it is apparent that the influence of rehearsal in task condition can be identified from the comparisons of Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3 (Table 4.3). As for the task condition of strategic planning, it could not be compared statistically, for it is included in each phase. However, the influence of strategic planning on CAF will be discussed later in the qualitative data analysis.

For the data analysis, the Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3, have been investigated by the parameter tests. These tests include the test of normality, paired t-test, and nonparametric test to answer the research question for the task condition of rehearsal and its influence on CAF.

On the other hand, the task characteristics are varied among Phase 1 and 2, Phase 2 and 3 as well as Phase 1 and 3 (Table 4.3). By comparing the CAF measures in Phase 1 and 2, Phase 2 and 3, the possible influence of familiarity of topic can be identified. To decide the effects of structure in speech-making, Phase 1 and Phase 3, Phase 2 and Phase 3 data have been analyzed.

Regarding the overall data analysis, the Phase 1, Phase 2, and Phase 3 recording data, have been run by the parameter test of normality, paired t-test, non-parametric tests for the results of task condition of rehearsal and the impacts on CAF. Meanwhile, the repeated measures ANOVA in SPSS has been analyzed for the results of task characteristics (familarity of information; structure) and their impacts on CAF. Next,

the rationale of the statitical analysis will be clarified.

Phase 2 Rehearsal and Phase 2, Phase 3 Rehearsal and Phase 3

Parameter Test

The parameter test refers to the statistical test of the average value and variance of parameters, which is an important part of inference statistics. When the population distribution is known (such as normal distribution), the statistical parameters of the population distribution can be inferred according to the sample data (Wang, 2012). To answer the research question whether the Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3 have any influence on students' language production in CAF, the parameter test of the paired t-test will be adopted. However, to run through the paired t-test, the data sets will be first tested for the normal distribution. If the test of normality is rejected for the data, the nonparametric test will be selected.

Test of Normality

To clarify the concept of test of normality, the definition of normal distribution will be first established. As a group, almost all results of human behavior are normally distributed. No matter what method is used to collect what kind of data, the distribution of large sample data is generally normal (Qin & Bi, 2015). Normal distribution is an important concept in statistical analysis because the rule of the parameter test needs to infer the results of the sample to the population, so it requires the data to be normally distributed (Qin & Bi, 2015). Therefore, the test of normality was conducted for Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3 to test for normal distribution. When the data are normally distributed, the parameter tests will be used. Otherwise, the non-parameter tests will be chosen.

Phase 1, Phase 2, and Phase 3 Data

ANOVA Analysis

The Analysis of Variance (ANOVA), also known as the "F-test", which is used to test

the significance of the mean difference between two or more samples (Qin & Bi, 2015). ANOVA is used to compare the variance between groups through dividing the variance between groups by the variance within groups, and then calculating the F value (Qin & Bi, 2015). The large F value can indicate that the difference between group variance is greater than the within group variance. If the F value reaches the significance level, then we can reject the null hypothesis (Qin & Bi, 2015). Once the null hypothesis is rejected, the alternative hypothesis is supported from the data in the research, indicating that the average value of each population is equal (Qin & Bi, 2015).

Why ANOVA?

The purpose of ANOVA is to test whether the variation of factors has influenced the dependent variables, which is used to test the main effects and interactions between variables (Kerr, Hall & Kozub, 2002). Its main function is to analyze the variation from different sources in the total variation of dependent variables, such as the variation caused by research and treatment (Qin & Bi, 2015). It is to determine the influence of controllable factors on the research results by analyzing the contribution of variation from different sources to the total variation. Specifically, the analysis of variance is used to test the significance of mean differences, classify the relevant factors, and estimate their effects on the total variation of dependent variables and the analysis of the interaction between factors (Qin & Bi, 2015). So, in the ANOVA analysis, the relationship between dependent and independent variables are examined.

Independent and Dependent Variables

The ANOVA analysis concentrates on the relationship between dependent variables and independent variables. Next, the concepts of variables, independent, and dependent variables will be identified.

To clarify, the term “variables”, in short, refers to things that may change.

Independent variable refers to the factors or conditions that researchers actively manipulate and cause the dependent variable to change, which is used for prediction. It is the variable selected by the researchers to study what effect they have on the dependent variable or what relationship they have with the dependent variable (Zhuo, 2009). The dependent variable, however, refers to the predicted variable. It is the variable that researchers use to observe the effect of changes with the independent variables. So, the independent variable is regarded as the "cause" while the dependent variable is viewed as the "effect" (Zhuo, 2009).

In this research, to answer the question "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", the independent variables are the manipulated interventions in task characteristics: familiarity of topic and structure as well as the task conditions: rehearsal and strategic planning in all 3 research phases. The dependent variables include the spoken language production of the above-mentioned indices of CAF measures, as they are observed to change in the 3 research phases.

If the data collected from the same subjects are under different conditions (three or more), repeated measures of ANOVA can be used for the parameter test (Qin & Bi, 2015). In this study, there are more than 3 conditions (3 research phases) being varied in task characteristics and task conditions. Therefore, the repeated measures of ANOVA should be adopted in the data analysis.

Repeated Measures of ANOVA

Repeated measures ANOVA refers to the measurement of an observation index of the same observation object at different time periods, which is used to analyze the changing patterns of the observation index at different times (Li, 2011). It is a design technique for repeated tests on the same dependent variable. The purpose is to study the significant differences among various treatments between subjects along with the

differences between several measurements (Liu & Liu, 2017). Through repeated measurements, the research subject is measured multiple times with multiple variables. In relation to this research, it is appropriate to apply the repeated measures ANOVA to the varied independent variables of task characteristics and task conditions and the multiple dependent variables of CAF measures (Kerr, Hall & Kozub, 2002).

Sphericity Assumption

In the analysis of variance between subjects, different processing data must be independent, that is, each subject is observed once, and the subjects are randomly assigned to different experimental conditions to meet the independent conditions. Then, the F-test can be correct (Qin & Bi, 2015). However, the repeated measurement design makes the repeated measurements on subjects under different conditions, which obviously violates the independence hypothesis of ANOVA between subjects and makes the traditional F-test lose its accuracy. Since the experimental design cannot meet the requirement of independence, it means that another hypothesis will be needed: the non-independence between different treatment levels, that is, the degree of dependence between different treatment levels must be roughly the same, which is called the sphericity assumption (Qin & Bi, 2015). The so-called sphere means that the difference between the values measured at different processing levels has equal variance. In other words, before we test whether there are significant differences between the treatment levels of each group, we first require that there is equal variance between them (Qin & Bi, 2015).

In SPSS, it will not only provide the results of repeated measures ANOVA, but also the results of Manchly's test of sphericity. If the significant value of Manchly's test of sphericity is greater than 0.05 ($P>0.05$), it means that the variances of the three variables are roughly equal, indicating that the data accept and satisfy the sphericity assumption. If it is less than 0.05, it indicates that the data does not meet the sphericity assumption. Then, the P value in the Multivariate Tests can be examined

(Qin & Bi, 2015).

Therefore, based on the SPSS analysis, the quantitative findings of task characteristics and task condition have been generated with the CAF spoken language production from different phases in this study.

Qualitative Data and Analysis

The qualitative data consist of two parts: online questionnaires and focus group interviews.

Online Questionnaires

The main purpose of conducting the online questionnaires is to get the larger sample' perceptions on different task characteristics and task conditions and their influence on spoken language production. The online questionnaires have been collected twice, once after phase 2, and once after phase 3 among the 3 classes. For this study, the semi-structured questionnaires are chosen with the mixture of closed and open-ended questions to present a series of questions. The students then responded and commented on them for their views on the research question. The semi-structured questionnaires share a clear structure, sequence, and focus, but the format is open-ended, enabling students to reply on their own terms (Cohen, Manion & Morrison, 2011). Therefore, rich data can be generated from the semi-structured questionnaires and findings can be found from the key words analysis of the online questionnaires.

In the questionnaires, questions were designed like "Which of the topics in phase 1, 2 and 3 are more familiar?"; "Can a more familiar topic help enhance your language production in CAF?". Meanwhile, matrix scale questions like "How do you think the following pre-task planning of strategic planning and rehearsal can help you prepare the speech" were devised to get the answers and later compare the results from previous research findings and triangulate with the quantitative data and focus group

interviews. For each question, open-ended questions were asked on the reasons why they chose the answers. With the key words analysis of the answers of the questionnaires, findings can be found to answer the research question in this study.

Focus Group Interviews

The purpose of conducting focus group interviews was to gather data, sample respondents' specific opinions, and evaluate or assess the influence of task characteristics and task conditions on the speech production in CAF from the students' viewpoints. Therefore, a focus group, semi-structured interview has been selected. When planning semi-structured interviews, a list of questions has been designed to explore answers to this study. During the interview, questions might not adhere to the plan. This flexible interview process can provide insight on students' view of the task characteristics and task conditions on the influence of spoken language production, in the learning process (Bryman, 2008).

Focus Group Participants

Altogether, there were 6 students for the focus group interviews. The interviewees were 2 students from each of the three classes. When selecting interviewees, the sample distribution characteristics such as different majors and gender were taken into consideration. Among the 6 participants, there were 50% of male and female students respectively, from the major of business management, computer science, and environmental science and engineering. Therefore, in the first focus group interview, all the 6 students participated. In the second one, there were 5 students being interviewed.

Collection of Focus Group Interviews

By preparing an interview outline in advance and organizing the focus group interviews, the first-hand raw materials of the interviews were obtained. Because of the COVID-19 safety protection, both focus group interviews were done through

online meetings in Chinese, as it was easier for the students to express their views freely. While interviewing, the students' real feelings, thoughts, and views of the research question were highly scrutinized. With the consent of the respondents, the interview data were recorded by the meetings, and then transcribed to form about 8000 words of written data, in Chinese. At the same time, the basic information of the respondents was encoded to hide their real names and avoid revealing personal privacy. All transcribed interview text materials were used in the study only after being approved by the respondents.

Focus Group Interviews Questions

The first focus group interview was after phase 1 and 2 of the study. Six students were invited with every two of them from one English class. To answer the first theme, task characteristics in the research question, two questions were asked on familiarity of information:

- For topic 1 and topic 2 in phase 1 and 2, which topic do you think you are more familiar with? Why?
- For the familiar topic, can it help to produce better language and why?

For the second theme, task conditions of strategic planning and rehearsal, several questions were asked.

- How do you think about your speech production when you are provided with strategic planning like: brainstorming for the speech contents; providing examples of model speeches and glossaries? Please explain why.
- How do you think about your speech production when you can rehearse the whole speech? Please explain why.
- Which helps you to produce language better? Strategic planning or rehearsal? Please explain why.

To probe for deeper insight, students were asked the reasons of their answers for further understandings of the research question.

The second focus group interview was after phase 1, 2, and 3 of the study. 5 students

from the previous group participated. To answer the first theme, task characteristics in the research question, questions were asked on familiarity of information:

- For the topic in phase 2 “describe a trip in Shenzhen” and phrase 3 topic “social software brings people closer or farther away”, which are more familiar to you? Please explain why.
- For the topic in phase 1 “shared bikes’ influence in life” and phrase 3 topic “social software brings people closer or farther away”, which are more familiar to you? Please explain why.
- For more familiar topics, can it help you speak more fluent, accurate, and complex spoken English? Why?

For degree of structure, questions were designed as follows.

- In the first two weeks, we did not design and explain the task structure of the speech: introduction, body, conclusion. In the third week, we explained the outline and structure of the speech. Can this help you speak more fluent, accurate, and complex spoken English? Why?

For the second theme, task conditions of strategic planning and rehearsal, several questions were asked.

- In the second week’s speech, we rehearsed the whole speech. In the third week’s task preparation, we made strategic planning in class, such as brainstorming the contents of the speech; provide examples of model speeches and vocabularies, and rehearsal. Compared with the second and third week, what kind of preparation is helpful for your oral speech in fluency, accuracy, and complexity? Where does each item help you? Please explain why.
- In the first week’s speech, we made strategic planning in class, such as brainstorming the content of the speech; provide examples of model speeches and vocabularies. In the third week, we made strategic planning in class, such as brainstorming the contents of the speech; provide examples of model speeches and vocabularies, and rehearsal. Compared with the first and third week, what kind of preparation is helpful for your speaking in fluency, accuracy, and complexity? Where does each item help you? Please explain why.

To get the overall answers to the research question, students were asked to compare all the three phrases of speeches for further understandings of the research question.

- Looking at the speech recordings in the first, second and third week, which week do you think the oral output was the most fluent, accurate, and complex? Please

explain why.

Data Analysis for Focus Group Interviews

With regards to the focus group interviews, the aim was to generate answers for the students' perceptions on task characteristics and task conditions and their influence on spoken language production. For the data analysis, the grounded theory has been adopted. The Grounded Theory Method (GTM) is described as a deductive induction of qualitative approach to research where theories are generated only from the investigation of data (Babbie, 2017). Next, the GTM as a qualitative method and the focus group data with the GTM analysis will be explained in detail.

The Grounded Theory as a Qualitative Method

The Grounded Theory is a qualitative research method, but not an entity of "theory". Its essence can be summarized as follows: the purpose of research is to generate theory, and theory must come from empirical data (Charmaz, 2006). For research, it is a process of systematically collecting and analyzing data from phenomena, discovering, developing, and testing theories from data. Consequently, the research result is a theoretical presentation of reality.

Traditional qualitative research itself does not include quantifications and may have many limitations, which is also one of the fundamental reasons for the emergence of the GTM (Wu, Wu & Ma, 2016). However, by the middle of the 20th century, the limitations of quantitative research had become increasingly prominent: (1) overemphasizing the representativeness of samples and lacking in-depth research on problems; (2) The complex and dynamic phenomenon of humanities and social sciences cannot be studied with quantitative methods such as statistics and measurement; (3) By verifying the assumptions based on existing theories, they can be made more refined to a certain extent, but new theories cannot be found (Wu, Wu & Ma, 2016). Just because the traditional qualitative and quantitative research have their own advantages and disadvantages, many scholars try to combine these two

methods: one is to introduce some methods in quantitative research (such as data analysis) into qualitative research; The other is to introduce some methods in qualitative research (such as the way of data collection) into quantitative research (Wu, Wu & Ma, 2016). The Grounded Theory belongs to the former. The quantitative analysis method is introduced into the qualitative research to overcome the contradiction between the insufficient depth and low validity in the quantitative research and the lack of standard procedures and poor reliability in the qualitative research (Wu, Wu & Ma, 2016).

Therefore, GTM, as a qualitative research method, adopts qualitative method and actively incorporates quantitative analysis in the process of data analysis. Through the combination of qualitative and quantitative techniques, GTM overcomes the basic contradiction between the insufficient depth and low validity, generalization, and reliability of quantitative research (He & Shi, 2009). In other words, GTM absorbs the advantages of quantitative research in qualitative research, using rigorous and systematic research procedures, and deductive induction in qualitative research to realize the "scientific" nature of research in a qualitative study (He & Shi, 2009). That is why the focus group data in Appendix 6 can include the quantified nodes in the analysis.

The Focus Group Data with the GTM Analysis

Analysis of focus group data uses GTM with the semi-structured interviews (He & Shi, 2009). The interview process is not based on pre-designed questions and fixed procedures, but just small talk between the interviewer and the respondents around the research question "whether the varied task characteristics and task conditions in the 3 research phases can influence students' spoken language production in CAF" to obtain in-depth, detailed, vivid and rich qualitative information (He & Shi, 2009). In addition, the semi-structured interviews can gather several respondents to talk at the same time (He & Shi, 2009). In general, the number of participants in the focus group

interview should be limited to 5-7. The participants should be the representatives, understand the situation, dare to speak, and have a common language with each other (He & Shi, 2009). The focus groups are the 6 representatives from 3 classes of English listening and speaking and the interviews have been done twice.

Analytic Procedure

The procedures of GTM for the focus group data include: 1) generating concepts from focus group data; 2) constantly comparing data and concepts and systematically asking generative theoretical questions related to concepts; 3) developing theoretical concepts and establishing the relationship between concepts; 4) constructing the theory, and striving to obtain the density, variability, and high degree of integration of theoretical concepts (Chen, 1999). In this study, the focus group interview data has been processed by the three levels of coding for the grounded theory (Chen, 1999).

Level 1 (Open Coding)

The first level coding (open coding) is an operational process of breaking up the collected focus group data, giving concepts, and then recombining them (Chen, 1999).

An example of the initial coding is as follows:

Initial Coding	Nodes
The “shared bikes” topic is better.	11
The understanding of the “shared bikes” topic is better.	3
They have experience of writing the topic of “shared bikes”.	2
They don’t have much experience of traveling in Shenzhen.	7
The topic “describe a trip in Shenzhen” is hard to express.	3
The topic “social networking” is more familiar.	7
The reasons of why “social networking” topic are more familiar	3
They have more understanding of the topic of “social networking” than the “shared bikes”.	2
The speech structure is helpful for the speech-making.	4
Providing glossaries in strategic planning is helpful.	7
Brainstorming in strategic planning is helpful.	6
Model speech in strategic planning is helpful.	5
Rehearsal is beneficial for speech-making.	6

Table 4.5 An Example of the Initial Coding

Level 2 (Axial Coding)

The main task of the second level of coding is to discover and establish various relationships among conceptual categories to show the organic relationship between various parts of the focus group data. These relationships can be causal, time sequenced, semantic, or situational (Chen, 1999).

An example of the secondary coding is as follows:

Secondary Coding	
The speech topics	The shared bikes topic
	The topic “describe a trip in Shenzhen”
	The social networking topic
Speech structure	
Strategic planning	
Rehearsal	
Familiarity of topic and spoken language production	
Speech structure and spoken language production	
Rehearsal and spoken language production	
Phase 1, 2, 3 task conditions and spoken language production	
Phase 1, 2, 3 spoken language production	

Table 4.6 An Example of the Secondary Coding

Level 3 (Selective Coding)

The Level 3 coding can follow these procedures: 1) to clarify the story line of the focus group data; 2) to describe the major and minor genera and their attributes and dimensions; 3) to test the established preliminary hypotheses and fill in the conceptual genera that need to be supplemented or developed; 4) to select the core conceptual genera; 5) to establish a systematic connection between the core genera and other genera (Chen, 1999). If more than one core genus can be found at the beginning of the

analysis, the related genera through continuous comparison will be connected, and those that are not closely related will be eliminated (Chen, 1999).

An example of the third level coding is as follows:
The Third Level Coding: Teaching conditions; Teaching actions; Teaching outcomes

After the coding and analysis of the focus group interviews, conclusions will be reached. Next, the analytic framework of this research study will be summarized.

4.4 ANALYTIC FRAMEWORK

For this research study, the quantitative and qualitative data have been analyzed as the framework shown in Figure 4.3.

For the quantitative data, the CAF measures from the convenience sampling groups were collected from the students' recordings. Accuracy was counted from the ratio of error-free clauses and the number of errors per hundred words in the speech recordings. Fluency was calculated from speech rate, number of pauses, repetitions, reformulation, and replacements in the speeches. Complexity was found from indices of lexical complexity and syntactic complexity in the speeches (Figure 4.3). All the data have been analyzed for the SPSS analysis to answer the research question on the task characteristics (familiarity of information, degree of structure) and task condition of rehearsal and their influence on spoken language production in CAF. Then, the qualitative data of online questionnaires have been summarized in the key words analysis. As for the focus group interviews, data has been processed, coded, and studied from the Grounded Theory Method. Both questionnaires and focus group interviews have provided insightful answers to the research question on task characteristics and task conditions and their influence of spoken language production. After that, both results from the quantitative and qualitative data have been compared

and triangulated to ensure the validity and reliability for the discussion and conclusion of the research study.

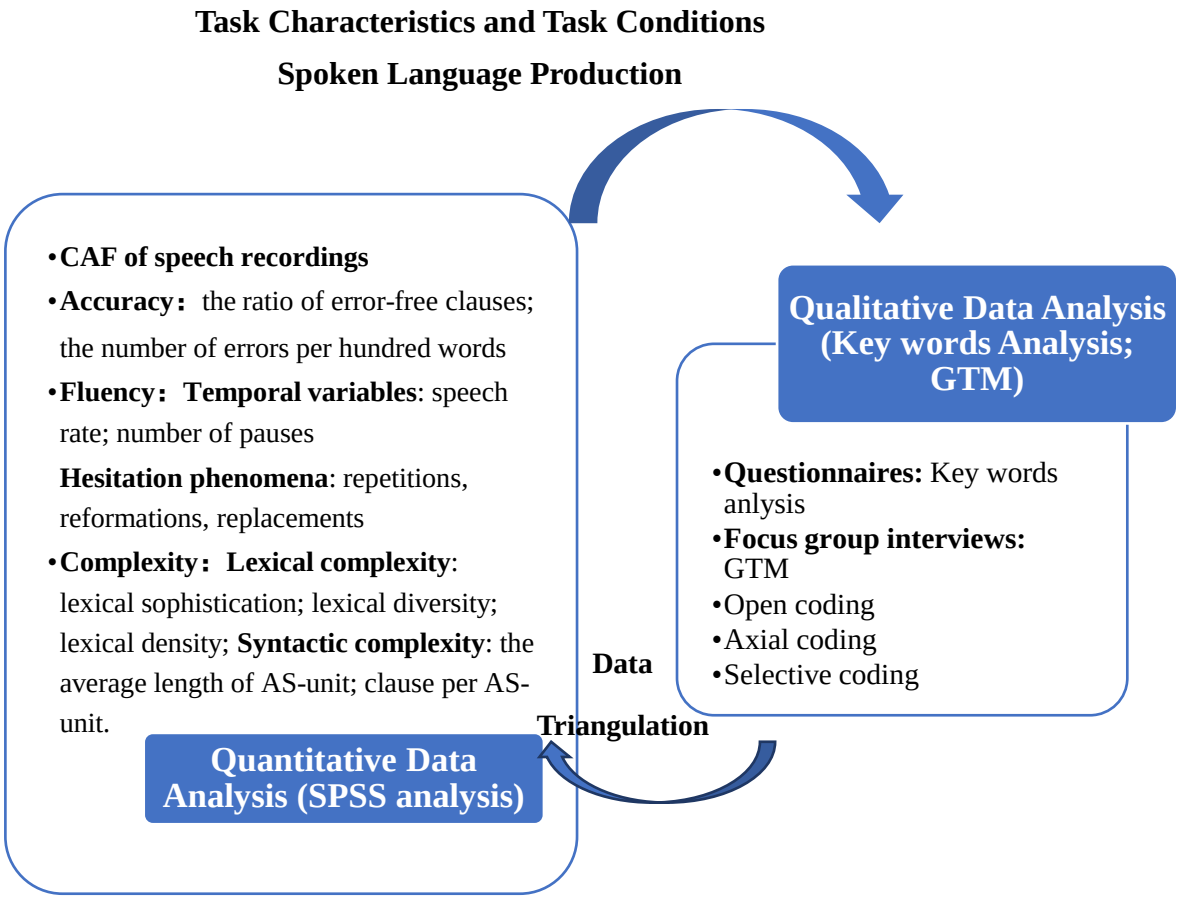


Figure 4.3: Analytic Framework of This Study

4.5 CHALLENGES FOR THE DATA ANALYSIS

The first challenge comes in the quantitative data analysis. As shown in the analytic framework (Figure 4.3), the measures of CAF can be complicated and the data analysis can be time-consuming. To cope with such a challenge, compromises were made for the CAF measures. For instance, the hesitation phenomena of false starts were omitted, because they were not distinct and significant in the recording transcriptions.

The second challenge lies in the time and plan of data analysis. The quantitative analysis of the oral recordings and the grounded theory analysis of focus group interviews were time-consuming. To ensure successful completion of data analysis, computer-aided tools such as SPSS and NVIVO were needed, which took time to learn. Therefore, thorough, and detailed planning were made to guarantee the accomplishment of data analysis.

4.6 CONCLUSION

This methodology chapter has included three principal sections. Philologically speaking, pragmaticism paradigms have been justified for this research, after exploring the themes of the paradigms in social sciences, the paradigm wars and pragmaticism. Methodologically, the mixed method research has been explicated from the definitions, purposes, and the design of the research. The explanatory triangulation model has been selected. Finally, the research design, research plan, data analysis, and analytic framework have been established in this study. The next chapter will proceed with a detailed investigation of data analysis.

Chapter 5 Data Analysis

As stated in the methodology chapter, this study has conducted mixed methods research to explore the task characteristics and task conditions and their influence on students' spoken language production. From the literature review in Chapter 3, the relationship between task characteristics and task conditions and their impacts on spoken language production can be summarized as follows: Tasks with concrete or familiar information can facilitate accuracy and fluency. Structured tasks can increase accuracy in the language production (Skehan, 2011). According to Ellis's (2013) hypothesis, rehearsal has no effect on CAF, but it is suggested that rehearsal is beneficial on the successive language performance of the same task (Ellis, 2005). In addition, strategic planning has shown positive effects on fluency and complexity.

To restate the research question, "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", this study analyzes both the quantitative data of students' recordings and qualitative data of questionnaires and focus group interviews upon the research question. Following the explanatory triangulation model of the mixed methods design, this data analysis chapter will be divided into three parts: quantitative data analysis, qualitative data analysis, and data triangulation. The quantitative data of students' recordings have been analyzed by the test of normality, paired t-test, and nonparametric tests for the influence of task condition of rehearsal in the students' speech production. For the task characteristics, the repeated measures of ANOVA have been adopted to examine the familiarity of topic and structure and their influence on spoken language production. Then, the qualitative data of questionnaires and the focus group interviews have been analyzed for the answers to the research question. Based on the findings, conclusion will be reached at the end of this chapter.

5.1 QUANTITATIVE DATA ANALYSIS

The quantitative data include speech recordings of the students for the measures of CAF in the 3 research phases (Table 5.1). For the quantitative data analysis, the CAF of the sample groups' recordings will be analyzed with the designed measures.

Task: speech making on a given topic			
Research phases	**Phase 1**	**Phase 2**	**Phase 3**
Task characteristics	Concrete: a familiar topic	Abstract: an unfamiliar topic	Concrete: a familiar topic
	Unstructured: without the speech outline	Unstructured: without the speech outline	Structured: with the speech outline
Task conditions	Strategic planning	Strategic planning & Rehearsal	Strategic planning & Rehearsal

Table 5.1: Three Research Phases

5.1.1 CAF Measures in Students' Recordings

As depicted in Chapter 4 Methodology, the CAF measures are shown in Table 5.2. Next, the specific CAF measures will be identified.

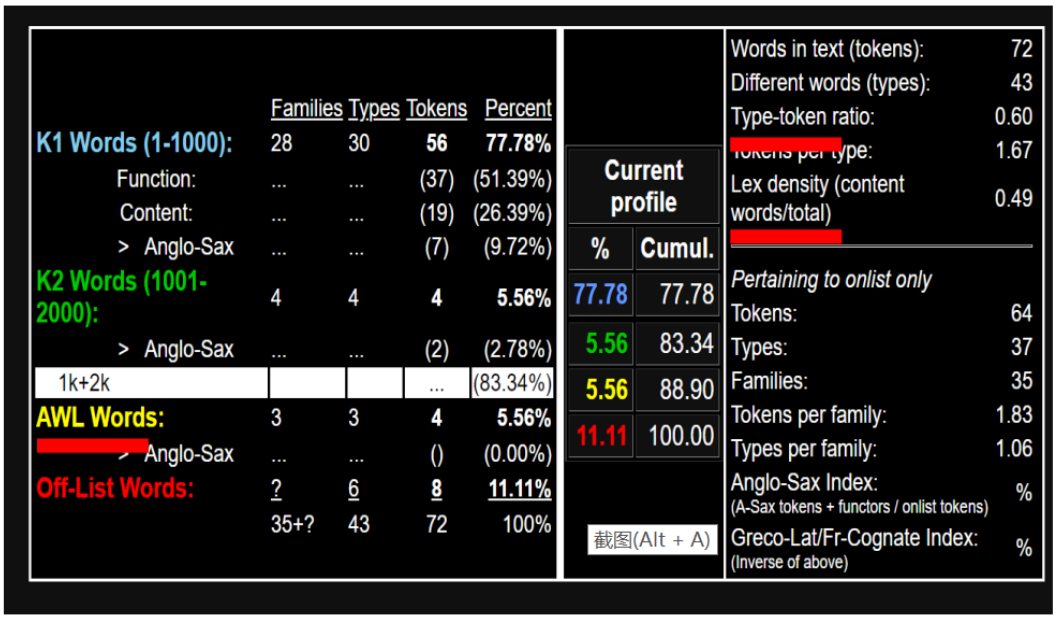
Complexity

Complexity contained two aspects: lexical and syntactic complexity. For lexical complexity, the measures were investigated in the lexical sophistication, diversity, and density (Table 5.2). Lexical sophistication was decided by the proportion of low-frequency words in the speeches. The AWL (Academic Word List) Beyond 2000 value of academic vocabulary represented the sophistication of vocabulary. For the lexical diversity, type-token ratio (TTR), which is the total number of different words

used (types) divided by the total number of words in the speech text (tokens), was investigated (Robinson, 1995). As for lexical density, the Content Word Ratio (CWR), which is the number of content words divided by the number of total words, was examined. For all the 3 indicators in lexical complexity, the online website VocabProfilers (https://www.lextutor.ca/vp/eng/) was adopted to retrieve the data of each recording. An example of the lexical complexity data can be seen in Table 5.3. The numbers show that the higher the indices of AWL words, TTR and CWR, the greater lexical complexity is presented.

Complexity	Lexical complexity: • Lexical sophistication: the proportion of low-frequency words in speaking • Lexical diversity: the ratio between parts of vocabulary (type) and token (TTR) • Lexical density: Content Word Ratio, CWR Syntactic complexity: • The average length of AS-unit • Clause per AS-unit
Accuracy	• The ratio of error-free clauses • The number of errors per hundred words
Fluency	Temporal variables: • Speech rate • The number of pauses and fillers Hesitation phenomena: • Repetitions • Reformations • Replacements

Table 5.2: The CAF Measures



	Families	Types	Tokens	Percent
K1 Words (1-1000):	28	30	56	77.78%
Function:	...	...	(37)	(51.39%)
Content:	...	...	(19)	(26.39%)
> Anglo-Sax	...	...	(7)	(9.72%)
K2 Words (1001-2000):	4	4	4	5.56%
> Anglo-Sax	...	...	(2)	(2.78%)
1k+2k			...	(83.34%)
AWL Words:	3	3	4	5.56%
> Anglo-Sax	...	...	()	(0.00%)
Off-List Words:	?	6	8	11.11%
	35+?	43	72	100%

Current profile	
%	Cumul.
77.78	77.78
5.56	83.34
5.56	88.90
11.11	100.00

截图(Alt + A)

Words in text (tokens):	72
Different words (types):	43
Type-token ratio:	0.60
[redacted]ype:	1.67
Lex density (content words/total)	0.49
Pertaining to onlist only	
Tokens:	64
Types:	37
Families:	35
Tokens per family:	1.83
Types per family:	1.06
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

Table 5.3: An Example of Lexical Complexity Data

For syntactic complexity, both "the average length of AS-unit" and "clause per AS-unit" were considered. First, the average length of T-unit was calculated by dividing the total number of words by the total number of T-units (Mu, 1988). The longer the average length of T-unit, the more complex the sentences are (Liu & Miu, 2018). As described in Chapter 4, the AS unit which involves "an independent clause or sub-clausal unit, together with any subordinate clause(s), associated with it from a single speaker's utterance" (Foster et al., 2000: 365) was used in the oral data analysis in this study. The average length of AS-unit was measured by the total words/the total number of AS-units. The longer the average length of AS unit, the more syntactic complexity is shown.

Second, the clause per T-unit was counted from the total number of clauses (C) divided by the total number of T-units (T). The higher C/T, the more clauses in the T unit, the more complex the sentences are (Xu, Zhang & Zhan, 2017). The clauses in each recording were determined through the syntactic complexity analyzer (https://aihaiyang.com/software/l2sca/single/). As discussed in Chapter 4, the measuring unit in this study was the AS-unit; So, the clause per AS-unit can be found

from C/AS; the number of clauses (C)/the total number of AS-unit (AS). The higher C/AS can indicate the more complex sentences in the speeches.

Accuracy

As shown in Table 5.2, the accuracy in the speech recordings was explored in two facets: the first was the ratio of error-free clauses, measured by the number of error-free clauses divided by the total number of independent clauses, sub-clausal units, and subordinate clauses multiplied by 100 (Foster & Skehan, 1996). It is suggested that the higher the ratio of error-free clauses, the more accurate the speech. The second was the number of errors per hundred words, which was calculated by the numbers of the errors divided by the total number of words and divided by 100 (Mehnert, 1998). The lower the number of errors per hundred words in the speeches, the higher the accuracy of the speeches.

Fluency

As described in Chapter 4, the fluency measure was examined in two layers: temporal variables and dysfluency elements of the hesitation phenomena. For the temporal variables, the speech rate and the number of pauses and fillers were counted. The speech rate was counted by the total words in the speech/minute. A higher speech rate can mean a more fluent speech. The numbers of pauses and fillers were detected from the transcripts of the total number of pauses and fillers, with features like pause and long pause: [P]; [LP]; and fillers: [ER]; [UH]; [URM] (Appendix 2). The number of pauses and fillers can indicate the degree the speaker withdraws from speaking as time is needed to plan the speech. The less time spent on planning the speaking with pauses and fillers, the more fluent the speech is (Ellis & Barkhuizen, 2005). The dysfluency elements of hesitation phenomena are related to the immediate and timely decisions to show achievable adjustments and improvements under the pressure of real-time communication (Skehan & Foster, 1999). To examine the hesitation phenomena, three aspects in the real-time speeches, repetitions, reformations, and

replacements, were scrutinized. The number of repetitions was marked by the immediate and verbatim repetition of a word or phrase in the speeches. The number of reformations was identified by phrases or clauses that were repeated with some modifications either to syntax, morphology, or word order. The number of replacements meant the number of lexical items that were substituted for another (Skehan & Foster, 1999). The lower the number of the 3 elements in the hesitation phenomena, the more fluency was displayed in the speeches.

The Indices and the CAF Measures

To summarize, the relationship of the indices and the CAF measures are shown in Table 5.4.

CAF Measures	Indices	The Relationship between the Indices and CAF Measures
Complexity	Lexical complexity: • lexical sophistication • lexical diversity • lexical density	 Lexical sophistication ↑ C↑ Lexical diversity ↑ C↑ Lexical density ↑ C↑
	Syntactic complexity: • the average length of AS-unit • clause per AS-unit	The average length of AS-unit ↑ C↑ Clause per AS-unit ↑ C↑
Accuracy	The ratio of error-free clauses	The ratio of error-free clauses↑ A↑
	The number of errors per hundred words	The number of errors per hundred words ↓ A↑
Fluency	Temporal variables: • speech rate; • number of pauses or fillers	 Speech rate ↑ F↑ Number of pauses or fillers ↓ F↑
	Hesitation phenomena: • repetitions • reformations • replacements	 Repetitions ↓ F↑ Reformations ↓ F↑ Replacements ↓ F↑

Table 5.4: The Relationship between the Indices and the CAF Measures

Data Processing

Next, the data processing of the speech recordings will be elucidated in the two procedures of transcription and calculation.

For the first step of the transcription, all the recordings from the sample groups for the 3 research phases were processed by Xunfei's Open Platform (https://www.iflyrec.com/) to gain initial transcription files of the 150 recordings. Then, each recording was listened approximately 4-5 times by the researcher to confirm on the pronunciation of words, phrases, and word order in the transcriptions (Appendix 2). After that, the fluency features of pauses, fillers, repetitions, reformations, and replacements were identified in each transcription. Each of the fluency features was double checked through careful listening.

After the transcription of 150 MP3 recordings, an SPSS spreadsheet was made based on the numbers of each index for the CAF measures. To start with, the lexical complexity indices were found from the above-mentioned website VocabProfilers. The recording transcription without fluency features was uploaded in the website to access data for lexical sophistication, diversity, and density. Then, the total number of words, the total number of clauses, and the total number of AS-units were counted to calculate the average length of AS-unit and the clause per AS-unit for syntactic complexity in each recording. Second, for accuracy, the numbers of error-free clauses and total numbers of independent clauses, sub-clausal units, subordinate clauses in each recording were calculated for the ratio of error-free clauses. Then, the number of the errors in each recording was spotted for the calculation of the number of errors per hundred words. Third, for fluency, the speech rate of each recording was found from the number of total words per minute. The pauses, fillers, repetitions, reformations, and replacements were traced from the fluency features in each recording of the transcriptions. A sample of the speech coding of CAF can be seen in Appendix 3. All the numbers in the SPSS spreadsheets were examined twice.

Data Analysis

Based on the above explanation of the relationship between the indices and the CAF measures (Table 5.4), the increase of lexical and syntactic complexity promote

complexity. The rise of the ratio of error-free clauses and the decrease of the number of errors per hundred words enhance accuracy. For fluency, the rise in speech rate and the decline of the numbers of pauses and fillers, repetitions, reformations, and replacements can develop fluency.

Therefore, by viewing the change of the indices in the 3 research phases, the analysis below explores the varied relationship of CAF measures and task characteristics and task condition to determine the answers for the research question. On one hand, the Phase 2 rehearsal data and Phase 2, and Phase 3 rehearsal data and Phase 3, have been investigated by the parameter tests including the test of normality, paired t-test, and the nonparametric test to answer the research question for the task condition of rehearsal and its influence on CAF. On the other hand, the Phase 1, 2, and 3 recording data, have been analyzed, by the repeated measures ANOVA in SPSS, for the results of task characteristics and their impacts on CAF.

Phase 2 Rehearsal and Phase 2, Phase 3 Rehearsal and Phase 3

As mentioned in Chapter 4, it is apparent that the influence of rehearsal in task condition can be identified from the comparisons of Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3 (Table 5.5). As for the task condition of strategic planning, it could not be compared statistically, for it is included in each phase. However, the influence of strategic planning on CAF will be discussed later in the qualitative data analysis.

As explained in Chapter 4, the test of normality will be conducted for Phase 2 rehearsal data and Phase 2, Phase 3 rehearsal data and Phase 3 to test for normal distribution. When the data are normally distributed, the parameter tests of the paired t-tests will be used. When they are not, the non-parameter tests will be chosen.

Task: speech making on a given topic				
Research phases	**Phase 2 rehearsal**	**Phase 2**	**Phase 3 rehearsal**	**Phase 3**
Task characteristics	Abstract: an unfamiliar topic	Abstract: an unfamiliar topic	Concrete: a familiar topic	Concrete: a familiar topic
	Unstructured: without the speech outline	Unstructured: without the speech outline	Structured: with the speech outline	Structured: with the speech outline
Task conditions	Strategic planning	Strategic planning & Rehearsal	Strategic planning	Strategic planning & Rehearsal

Table 5.5: Phase 2 and Phase 3 Research Design

Phase 2 Rehearsal and Phase 2 (P2): Test of Normality

For the test of normality, the values in the W-test will be examined. The W-test was proposed by S. S. Shapiro, M. B. Wilk in 1965 (Liu & Hu, 2005). It is suitable for the normality test when the sample size is $n \leq 50$; As this study contains 30 samples, the results in the W-test will be reviewed. When the P value of both P2 Rehearsal and P2 is over 0.05 ($P>0.05$), the test of normality is accepted (Liu & Hu, 2005).

Shapiro-Wilk (P Value)-	P2 Rehearsal	P2
Lexical sophistication	0.000	0.000
Lexical diversity	0.005	0.059
Lexical density	0.059	0.012
Syntactic complexity: average length of AS unit	0.973	0.706
Syntactic complexity: clause per AS unit	0.208	0.005
Accuracy: the ratio of error-free clauses	0.311	0.027
Accuracy: the number of errors per 100 words	0.000	0.581
Fluency: speech rate	0.923	0.555
Fluency: pauses and fillers	0.003	0.004
Fluency: repetitions	0.000	0.000
Fluency: reformations	0.144	0.003
Fluency: replacements	0.000	0.000

Table 5.6: Test of Normality of P2 Rehearsal and P2

In this study, when both the $P>0.05$ in Phase 2 rehearsal and Phase 2, the tests of

normality are accepted. So, only two indices: the syntactic complexity of the average length of AS unit and the fluency of speech rate (P>0.05) are tested for normality (Table 5.6). Therefore, these two can be tested for the paired t-test. For the rest which are not tested for normality, the non-parameter test will be conducted (Qin & Bi, 2015).

Paired T-test

The paired sample t-test is also known as non-independent sample test. It is a repeated measurement design test, whose function is to compare the average values of two groups. The purpose is to test whether the difference between average values reaches the significant level. The two groups of data can be from the same subjects and used to compare the results before and after the experiment (Qin & Bi, 2015). In this study, the subjects are the 30 students before and after the speech rehearsal. The paired t-tests will be compared to view whether there are significant levels before and after the rehearsal.

In the below paired sample t-test, P<0.05 for both syntactic complexity of average length of AS unit and fluency of speech rate, which means both samples are significantly different. As for the mean value, both P2>P2 rehearsal (6.98>6.40; 98.47>85.93) (Table 5.7). Therefore, the syntactic complexity of average length of AS unit and the fluency of speech rate in Phase 2 (after the rehearsal) is better than P2 rehearsal (before the rehearsal). Thus, rehearsal has a promoting influence on these two indices in Phase 2.

		Paired Sample Test Sig.(2-tailed)	Paired sample statistics Mean
Syntactic complexity of average length of AS unit	P2 rehearsal	0.010	6.40
	P2		6.98
Fluency of speech rate	P2 rehearsal	0.000	85.03
	P2		98.47

Table 5.7: Paired T-test for P2 Rehearsal and P2

Nonparametric Test

The nonparametric test does not consider the parameters and distribution pattern of the population, but obtains information from the sample itself to test the hypothesis of the distribution or distribution position of the population it represents (Yang & Xie, 2014). Because this kind of method is not limited by the population parameters, it verifies the distribution rather than parameters. So, it is called the nonparametric test, also known as the arbitrary distribution test or free distribution test (Yang & Xie, 2014).

Mann-Whitney U-test

Provided that there is to be a comparison of two groups, but the data violates the normal distribution test conditions, the non-parametric test, also called the Mann-Whitney U-test, can be applied. The Mann-Whitney U-test is a sum of rank test, that is, the rank comparison. Its null hypothesis is that the rank sum of two groups is the same (Qin & Bi, 2015). This method combines the data of two independent samples and sorts them from small to large. The rank of the smallest observation is 1, followed by 2, and so on. If there are the same observations, the average of their ranks is used for sorting (Qin & Bi, 2015). Therefore, the Mann-Whitney U-test will be adopted to compare the results of the non-parameter test of the indices in Phase 2 rehearsal and Phase 2.

Mann-Whitney U-test: Asymp. Sig.(2-tailed) P2 Rehearsal and P2	
Lexical sophistication	0.555
Lexical diversity	0.744
Lexical density	0.055
Syntactic complexity: clause per AS unit	0.286
Accuracy: the ratio of error-free clauses	0.739
Accuracy: the number of errors per 100 words	0.574
Fluency: pauses and fillers	0.022
Fluency: repetitions	0.560
Fluency: reformations	0.221
Fluency: replacements	0.826

Table 5.8: Mann-Whitney U-test Results of Phase 2 Rehearsal and Phase 2

For most of the indices, their Mann-Whitney U-test assymp. Sig (2-tailed) P>0.05. The significant difference is rejected when P>0.05 (Table 5.8). Phase 2 rehearsal and Phase 2 are almost the same for complexity, accuracy, and most of the fluency factors. However, the fluency index of pauses and fillers (P=0.022<0.05) is significantly different in Phase 2 rehearsal and Phase 2. The mean rank of Phase 2 rehearsal is 35.62, while Phase 2 is 25.38. This indicates that pauses and fillers are showing more in Phase 2 before rehearsal than Phase 2 after the rehearsal, which could mean rehearsals are likely to decrease pauses and fillers in speech-making in the Phase 2 research.

To summarize the results in Phase 2 rehearsal and Phase 2 in the same task, rehearsal could promote the 2 indices in syntactic complexity of average length of AS unit and the fluency of speech rate. Meanwhile, rehearsal can reduce the pauses and fillers to enhance fluency.

Phase 3 Rehearsal and Phase 3 (P3)

Likewise, in Phase 3 research, rehearsal is also applied before and after the speech-making. Next, the similar approaches of data analysis and results will be reported. The Test of Normality, in Phase 3 Rehearsal and Phase 3, is shown in Table 5.9.

Shapiro-Wilk (P Value)-	P3 Rehearsal	P3
Lexical sophistication	0.135	0.341
Lexical diversity	0.260	0.497
Lexical density	0.823	0.985
Syntactic complexity: average length of AS unit	0.243	0.181
Syntactic complexity: clause per AS unit	0.097	0.331
Accuracy: the ratio of error-free clauses	0.541	0.510
Accuracy: the number of errors per 100 words	0.234	0.495
Fluency: speech rate	0.391	0.298
Fluency: pauses and fillers	0.002	0.000
Fluency: repetitions	0.000	0.000
Fluency: reformations	0.510	0.015
Fluency: replacements	0.000	0.000

Table 5.9: Test of Normality of P3 Rehearsal and P3

In Phase 3 research, the lexical and syntactic complexity indices, the accuracy indices, and the fluency of speech rate are accepted for the test of normality (P>0.05). Thus, these indices will be analyzed by the paired t-test. As for the fluency of pauses and fillers, repetitions, reformations and replacements, the P values of both the Phase 3 rehearsal and Phase 3 are not larger than 0.05 (P<0.05). So, these indices will be evaluated through the nonparametric test.

Paired T-test

As is known from the above analysis, 8 indices from the CAF measures in the Phase 3 rehearsal and Phase 3 are tested for the normality. Thus, they are acceptable to run the paired t-test. The results are demonstrated in Table 5.10.

		Paired Sample Test Sig.(2-tailed)	Paired Sample Statistics Mean
Lexical sophistication	P3 rehearsal	0.125	
	P3		
Lexical diversity	P3 rehearsal	0.452	
	P3		
Lexical density	P3 rehearsal	0.320	
	P3		
Syntactic complexity of average length of AS unit	P3 rehearsal	0.097	
	P3		
Syntactic complexity: clause per AS unit	P3 rehearsal	0.564	
	P3		
Accuracy: the ratio of error-free clauses	P3 rehearsal	0.200	
	P3		
Accuracy: the number of errors per 100 words	P3 rehearsal		
	P3	0.799	
Fluency of speech rate	P3 rehearsal	0.000	81.52
	P3		90.97

Table 5.10: Paired T-test for P3 Rehearsal and P3

From the above results, all the complexity and accuracy measures indices are not significantly different (P>0.05), which means that when the indices between Phase 3 rehearsal and Phase 3 are compared, there is not much difference. Only one index, the fluency of speech rate (P=0.000<0.05) means that the two groups are different in

significance. By viewing the paired sample statistics mean, the mean value before rehearsal in Phase 3 is 81.52. After rehearsal, it increases to 90.97, which indicates that rehearsal can increase the speech rate in fluency.

Non-parametric Mann-Whitney U-test

The 4 fluency indices in Phase 3 that are rejected in the test of normality are assessed by the non-parametric Mann-Whitney U-test. The results are displayed in Table 5.11. The results in Phase 3 research indicate that all the 4 fluency indices are not statistically significant (P>0.05). Therefore, there is no interest to compare the results of both groups for the 4 fluency indices.

Mann-Whitney U-test: Asymp. Sig.(2-tailed)	Fluency: pauses and fillers	Fluency: repetitions	Fluency: reformations	Fluency: replacements
P3 rehearsal P3	0.187	0.358	0.223	0.333

Table 5.11: Mann-Whitney U-test Results of Phase 3 Rehearsal and Phase 3

Summary of Rehearsal in Phase 2 and Phase 3 Research

Table 5.12 summarizes the findings of rehearsal in Phase 2 and Phase 3 research. They show that the fluency of speech rate is likely to increase with the influence of rehearsal. As for the effects of rehearsal on syntactic complexity and the rest of the indices in fluency, further research is needed.

Rehearsal	Phase 2 research: 1. Syntactic complexity-average length of AS unit ↑ 2. Fluency-speech rate ↑ 3. Fluency-pauses and fillers ↑
	Phase 3 research: Fluency-speech rate ↑

Table 5.12: Summary of Findings in Phase 2 and Phase 3 Research

Phase 1, 2, 3 Data Analysis

Looking back at the 3 research phases in Table 5.1, the task characteristics are varied

among Phase 1 and 2, Phase 2 and 3 as well as Phase 1 and 3. By comparing the CAF measures in Phase 1 and 2, Phase 2 and 3, the possible influence of familiarity of topic can be identified. To decide the effects of structure in speech-making, Phase 1 and Phase 3, Phase 2 and Phase 3 data have been analyzed. Respecting the analysis of task characteristics in the research question, the repeated measures of ANOVA has been chosen as mentioned in Chapter 4.

Phase 1, 2, and 3 Manchly's Test of Sphericity

Table 5.13 below shows the P values of the CAF indices in Phase 1, 2, and 3 for Manchly's Test of Sphericity (sample analysis see Appendix 4). As it is demonstrated in Table 5.13, on one hand, the complexity, accuracy, and the fluency indices of speech rate, repetitions and reformations are accepted for the Manchly's test of sphericity ($P>0.05$). They are viewed as statistically significant for further discussion of the Phase 1, 2 and 3 research. Next, the sphericity assumed P value in Tests of Within-Subjects Effects will be explored. As displayed in Table 5.13, the three lexical complexity indices, the syntactic complexity of clause per AS unit, the accuracy index of the number of errors per 100 words, fluency indices of speech rate, and reformations ($P<0.05$), are significant statistics. The syntactic complexity of average length of AS unit, accuracy of the ratio of error-free clauses, and fluency of repetitions ($P>0.05$), are not statistically significant and they will not be included in the later discussion.

On the other hand, the fluency indices of pauses, fillers, and replacements are rejected for Manchly's test of sphericity. Then, the Wilks' Lambda P value in the Multivariate Tests will be investigated. For the fluency index of pauses and fillers ($P=0.010<0.05$), it is statistically significant for further comparison of this index in the Phase 1, 2, and 3 research. For the fluency index of replacements ($P=0.106>0.05$), it is not significant in statistics. So, this index will not be examined.

Manchly's Test of Sphericity (P Value)-		Tests of Within-Subjects Effects (sphericity assumed P value)	Multivariate Tests (Wilks' Lambda P value)
Lexical sophistication	0.674	0.000	
Lexical diversity	0.706	0.000	
Lexical density	0.182	0.027	
Syntactic complexity: average length of AS unit	0.491	0.331	
Syntactic complexity: clause per AS unit	0.720	0.046	
Accuracy: the ratio of error-free clauses	0.716	0.526	
Accuracy: the number of errors per 100 words	0.370	0.002	
Fluency: speech rate	0.344	0.014	
Fluency: pauses and fillers	0.001		0.010
Fluency: repetitions	0.417	0.119	
Fluency: reformations	0.065	0.002	
Fluency: replacements	0.000		0.106

Table 5.13: Phase 1, Phase 2, and Phase 3 Manchly's Test of Sphericity

Phase 1 (P1), Phase 2 (P2), and Phase 3 (P3) Results

From the above analysis, the dependent variables are narrowed down to the below 8 indices, which are statically significant between groups. To answer the research question for the task characteristics (familiar of information; structure) and their influence on CAF, the Phase 1, 2, and 3 results of these 8 indices of the pair comparisons of within group data will be manifested in Table 5.14.

For the lexical sophistication, the P value of Phase 1, 2, and 3 are statistically significant (P=0.000<0.05). From the I-J value, Phase 3 (I)-Phase 1 (J)=0.015; Phase 3 (I)-Phase 2 (J)=0.039; Phase 1 (I)-Phase 2 (J)=0.023. It is obvious to see Phase 3> Phase 1>Phase 2. By looking back at the task characteristics interventions of the three research phases in Table 5.1, Phase 3 includes familiar topic and structure, Phase 1 familiar topic and unstructured, and Phase 2 is an abstract topic and unstructured task. A possible conclusion of task characteristics can be drawn that the more familiar topic

and the structure of speech-making could foster lexical sophistication.

Pairwise Comparisons		P 1 (I)		P 2 (I)		P 3 (I)	
		P 2 (J)	P3 (J)	P1 (J)	P3 (J)	P 1 (J)	P 2 (J)
Lexical sophistication	Mean difference (I-J)	0.023	-0.015	-0.023	-0.039	0.015	0.039
	Sig.	0.000	0.000	0.000	0.000	0.000	0.000
Lexical diversity	Mean difference(I-J)	-0.001	0.066	0.001	0.067	-0.066	-0.067
	Sig.	0.952	0.000	0.952	0.000	0.000	0.000
Lexical density	Mean difference(I-J)	0.018	-0.013	-0.018	-0.031	0.013	0.031
	Sig.	0.183	0.189	0.183	0.006	0.189	0.006
Syntactic complexity: clause per AS unit	Mean difference(I-J)	-0.078	-0.068	0.078	0.010	0.068	-0.010
	Sig.	0.025	0.027	0.025	0.739	0.027	0.739
Accuracy: the number of errors per 100 words	Mean difference(I-J)	0.000	0.000	0.000	0.000	0.000	0.000
	Sig.	0.003	0.091	0.003	0.040	0.091	0.040
Fluency: speech rate	Mean difference(I-J)	-10.667	-3.167	10.667	7.500	3.167	-7.500
	Sig.	0.012	0.321	0.012	0.051	0.321	0.051
Fluency: pauses and fillers	Mean difference(I-J)	1.667	-1.200	-1.667	-2.867	1.200	2.867
	Sig.	0.006	0.233	0.006	0.012	0.233	0.012
Fluency: reformations	Mean difference(I-J)	0.533	-2.200	-0.533	-2.733	2.200	2.733
	Sig.	0.377	0.017	0.377	0.003	0.017	0.003

Table 5.14: Phase 1, Phase 2, and Phase 3 Pairwise Comparisons

For lexical diversity, the P values of Phase 1 and Phase 3, Phase 2 and Phase 3 are statistically significant ($P<0.05$). To see the I-J value, Phase 1 (I)-Phase 3 (J) =0.066, while Phase 2 (I)-Phase 3 (J)= 0.067, which means that the lexical diversity in both

Phase 1 and Phase 2 are better than Phase 3. In this case, the familiar topic and the structure are not beneficial for lexical diversity.

With respect to lexical density, only the P value of Phase 3 and Phase 2 shows statistical significance (P=0.006<0.05). From the I-J value, Phase 3 (I)-Phase 2 (J)= 0.031, which shows Phase 3>Phase 2. Between Phase 3 and Phase 2, Phase 3 involves familiar topic and structure, while Phase 2 has not. Therefore, the lexical density could be promoted from the task characteristics of familiarity of topic and structure.

Concerning the syntactic complexity of clause per AS unit, the P value of Phase 2 and Phase 1, Phase 3 and Phase 1 are statistically significant (P<0.05). From the I-J value, Phase 2 (I)-Phase 1 (J)=0.078, Phase 2>Phase 1, which could not prove that the familiar topic could enhance syntactic complexity of clause per AS unit. However, for the I-J value, Phase 3 (I)-Phase 1 (J)=0.068, Phase 3 is better than Phase 1. In the Phase 3 and Phase 1, the only different intervention in task characteristic is structure. Thus, it is indicated that the more structured speech task can encourage syntactic complexity of clause per AS unit.

For the accuracy of the number of errors per 100 words, the P value of Phase 2 and Phase 1, Phase 2 and Phase 3 are statistically significant (P<0.05). However, both the I-J value of Phase 2 and Phase 1, Phase 2 and Phase 3 are 0.000, which does not reveal a difference in Phase 2 and Phase 1, Phase 2 and Phase 3.

Regarding the fluency of speech rate, the P value of Phase 2 and Phase 1 are statistically significant (P<0.05). From the I-J value (10.667), Phase 2 is better than Phase 1, which cannot explain that the familiar topic and structure increase speech rate.

Moving to the fluency of pauses and fillers, the P value of Phase 1 and Phase 2 are

statistically significant (P=0.006<0.05). From the I-J value, Phase 1 (I)-Phase 2 (J)=1.677. It is shown that Phase 1 has more pauses and fillers, which means less fluency than Phase 2. However, Phase 1 has a more familiar topic but with more pauses and fillers than Phase 2, which cannot specify that the familiar topic promotes fluency. Possible explanation for the better fluency result of pauses and fillers in Phase 2 may be about the rehearsal, which could reduce the fluency of pauses and fillers.

For the fluency of reformations, the P value of Phase 3 and Phase 1, Phase 3 and Phase 2 are statistically significant (P<0.05). From the I-J value, Phase 3 (I)-Phase 1 (J)=2.200, Phase 3 (I)-Phase 2 (J)=2.733. These indicate that Phase 3 has more reformations and less fluency than Phase 1 and Phase 2, which cannot satisfy the conclusion that the task characteristics of familiar topic and structure enhance fluency. Possible explanation for the better fluency results in Phase 1 could be that the topic is relatively easier than the one in Phase 3. In the meantime, the 2-minute speech task in Phase 3 is comparatively more demanding in the limited attentional resources of the learners to organize the language than in Phase 1 and Phase 2, with the time length of 1 minute.

Summary of Task Characteristics in Phase 1, 2, and 3 Research

Table 5.15 below can summarize the findings of task characteristics (familiarity of topic; structure) on spoken language production in Phase 1, 2, and 3 research.

Combining the results of the task condition of rehearsal in Phase 1, 2, and 3 research, the quantitative answer for the research question in this study can be summarized in Table 5.15. The more familiar topic and structure could promote lexical sophistication and density. Also, the structured task could be beneficial for the syntactic complexity of clause per AS unit. Considering the task condition, the fluency of speech rate is likely to increase with the influence of rehearsal.

Task characteristics	Familiarity of information ↑; Structure ↑	Lexical sophistication ↑
	Familiarity of information ↑; Structure ↑	Lexical density ↑
	Structure ↑	Syntactic complexity of clause per AS unit ↑
Task condition	Rehearsal ↑	Fluency of speech rate ↑

Table 5.15: Summary of Findings in Phase 1, 2, and 3 Research

5.2 QUALITATIVE DATA ANALYSIS

The qualitative data consists of two parts: questionnaires and focus group interviews. Both parts of the data will later be compared with the above quantitative findings for data triangulation in this study. First, the qualitative data of questionnaires will be reported.

5.2.1 Online Questionnaires

The online questionnaires have been designed to get students' views on task characteristics and task conditions and their influence on spoken language production. Data has been collected twice, once after Phase 2 and once after Phase 3. Both questionnaires have attained the qualitative opinions from more than 100 students in the 3 classes to serve as the majority's views on the research question for data triangulation. The sample questionnaire data can be seen in Appendix 5.

First Online Questionnaire (Phase 1 and Phase 2)

In the first questionnaire, 114 answers were collected for students' opinions on Phase 1 and Phase 2 research. Among the 114 students, about 75% were male and 25% female, with their majors ranging from computer science to architecture, automation, business management, energy power, civil engineering, and environmental engineering.

Viewing the task characteristic of familiarity of topic, about 57% of them considered the topic “shared bikes” in Phase 1 was more familiar than the Phase 2 topic “describe a trip in Shenzhen”, while 42% of them thought the opposite. Some of the students explained that they were new in Shenzhen at that time so that there was not much time to go out in Shenzhen, while the “shared bikes” topic was more related to their daily life. So, it was easier to speak about the topic of “shared bikes”.

Moving to the task conditions in Phase 1 and Phase 2, rehearsal received the most support with 77.19% of the students thinking it helpful for speech preparation. By looking at the keyword analysis of rehearsal, most students agreed that it gave them some time and opportunity to prepare. Meanwhile, the speech could be more fluent with rehearsal and have fewer tense feelings. Turning to strategic planning, providing glossaries, model speeches, and brainstorming have taken about 54%, 46% and 38% respectively. When the glossaries were provided, most students explained that knowing the vocabulary items of the speech could provide clearer logic regarding what to say. In the meantime, with the help of model speeches, logics, examples, and inspirations were what the students considered beneficial. Regarding brainstorming, they claimed that it was a good chance to compare notes with their classmates and get to know other viewpoints. The participants ranked the task conditions to show their views on whether they were helpful in the task preparation. For rehearsal, 87.72% of them regarded it most helpful. For glossaries and model speeches, about 69.3% and 62.3% respectively of them acknowledged it as most helpful. However, for brainstorming, about 32.46% thought it most helpful.

To evaluate themselves in the Phase 1 and Phase 2 language production in CAF, 51.75% of the students presumed that Phase 2 was better than Phase 1, while 48.25% of them believed Phase 1 better than Phase 2. For the supporters of Phase 2, the most frequent words appearing in their views were rehearsal, time, and preparation. So, it could be concluded that rehearsal has some promoting influence on spoken language

production. While supporters of Phase 1 claimed that the topic was closer and more familiar in their life. Others thought that providing glossaries and model speeches helped in speech preparation. Therefore, the familiar topic and strategic planning could enhance speech production as well.

Second Online Questionnaire (Phase 1, 2, and Phase 3)

In the second questionnaire, 112 answers were collected for the students' opinions on the Phase 1, 2, and 3 research. Among the 112 students, about 77 % were male and 22% female, with the same majors as the first time.

The task characteristic of familiarity of topic in Phase 2 and Phase 3, has been asked which is more familiar. Among the 112 students, 72 of them, which was about 64% considered that Phase 3 topic "social networking" was more familiar than the Phase 2 topic "describe a trip in Shenzhen". By viewing the keywords analysis, the most frequent quoted reasons were "more familiar and common topic in Phase 3" and "structure is presented in Phase 3". However, 40 out of 112 students, which was 35.71%, voted in favor of Phase 2 topic familiarity with most reasons of "personal experience and simpler speech" in the keyword analysis.

Additionally, Phase 1 and Phase 3 have been compared for familiarity of topic. Among the 112 students, 73 of them, which was about 65% recognized that Phase 3 topic "social networking" was more familiar than Phase 1 topic "shared bikes". While about 39 out of 112, which was about 34.82% of them reflected that the Phase 1 topic more familiar.

When asking the question "Can a more familiar topic help you produce more fluent, accurate, and complex language? Please explain why.", a very high proportion of the students, 83.93%, which was 94 out of 112, perceived that it was very helpful with the more familiar topic. Considering the keyword analysis, most of them held the views

that “the more familiar topic can help generate the opinions for spoken language production”.

Concerning task characteristic of structure, there was no outline and structure of the speech in Phase 1 and Phase 2. In Phase 3, the explanation and preparation of the outline and structure of the oral speech was added. Students were asked to compare the oral output in Phase 1, 2, and 3 to see whether the outline and structure of the speech in Phase 3 helped them produce more fluent, accurate, and complex language. Examining the answers to this question, 88 out of 112, which was 78.57%, agreed that the Phase 3 speech structure was very helpful for language production. The most frequently quoted reasons were that “the speech structure can clear the logic and contents in speech-making”.

With regards to the task conditions, Phase 1 and Phase 3 are compared. In Phase 1, there were only strategic planning of paired discussions of brainstorming, providing model speeches and glossaries. In Phase 3, both strategic planning and rehearsal were applied. 99 out of 112 students, which was 88.39%, considered that Phase 3 task conditions were more helpful to prepare the oral speech and express more fluent, accurate, and complex language than Phase 1. The most frequently mentioned reasons came with rehearsal in Phase 3. A sample of the key words analysis of open question can be found in Appendix 6.

As for the last question, it was asked that “from Phase 1, 2 and 3, which one do you think your oral output is the most fluent, accurate, and complex? Please explain why.” About 58% supported Phase 3 with the reasons of “more familiar topic and having more time to prepare”. Compared with Phase 3, 25% of the students, which was less than half of the supporters of Phase 3, considered Phase 2 language production better, claiming that it was more related to their personal experience. The least supporters came in Phase 1, about 16% of them believed that it was better in spoken language

production.

To summarize results in the online questionnaires, the task characteristics of familiarity of topic and structure and the task conditions of strategic planning and rehearsal in the general participant's viewpoints are deemed beneficial for students' spoken language production in CAF.

5.2.2 Focus Group Interviews

After Phase 2 and Phase 3, two focus group interviews, one about 20 minutes and the other about 30 minutes, were conducted to 6 representatives from the 3 classes, for their specific viewpoints on the task characteristics and task conditions and their influence on spoken language production. A sample of focus group interview questions can be found in Appendix 7. For the data triangulations of this study, the interviews will provide insightful qualitative data from the focus group regarding the research question. Next, the interviews have been analyzed by the grounded theory.

Data Analysis and Results

Based on the grounded theory and following the corresponding operating procedures of NVivo 12 plus, this study analyzed the collected interview text data, conceptualized the phenomena through the three-level coding, namely the initial coding, secondary coding and the third level coding (Chen, 2015). Then, the concepts were integrated into categories, and later generated into the core categories, to seek the relevant factors in the interviews (Chen, 2015). Finally, the answers to the research question of the influence of task characteristics and task conditions on students' spoken language production were revealed.

Consistency Test

In the interest of the reliability and consistency of the coding of focus group

interviews, two coders (the researcher and a colleague of the researcher) have done the initial coding of the interview data in NVivo 12 plus operations. Based on the Cohen's Kappa (Cohen, 1960), the interrater reliability can be calculated. If the reliability is greater than 0.8, it means that the reliability is acceptable. Classifying the two coders, one is coder A, and the other is coder B. The specific operation process is as follows: (1) train coder B to fully understand the research purpose, category meaning and coding operation requirements. (2) The text is then coded by coder B to compare the results with coder A. (3) The same coding results of both coders are counted, and then the reliability of the category system in this study is calculated and judged according to the following formulas.

$$K_{AB} = \frac{2M_{AB}}{N_A + N_B}$$

In the above formula, M_{AB} is the same number of the codes from both coders. N_A and N_B are the numbers of codes from each coder. K_{AB} is the average mutual agreement between the coders.

$$R = \frac{n \times K}{1 + (n - 1) \times K}$$

When K_{AB} is calculated, the discriminant reliability R can be counted. In this formula, n is the number of coders, and K is the average mutual agreement result from the previous formula. Therefore, the average mutual agreement degree of the two coders is calculated as follows: the number of identical codes in this study is 51. N_A and N_B represent the number of codes of two coders, 54 and 56 respectively. After calculation, K_{AB} is 0.93, and the reliability discrimination formula is brought in to calculate R (R=0.96), which is greater than 0.8, so, it passes the consistency test.

Initial Coding (Open Login)

In the initial coding of interview data, it is proposed to let the original data "speak" for itself. The concept categories are found from the text data. The keywords and conceptualization are extracted (Chen, 1999). The detailed analysis of themes and nodes can be found in Appendix 10. Each line of the focus group interviews is coded with the initial coding. The most common nodes of the initial coding are presented in Table 5.16.

Initial Coding	Nodes
The "shared bikes" topic is better.	11
The understanding of the "shared bikes" topic is better.	3
They have experience of writing the topic of "shared bikes".	2
They don't have much experience of traveling in Shenzhen.	7
The topic "describe a trip in Shenzhen" is hard to express.	3
The topic "social networking" is more familiar.	7
The reasons of why "social networking" topic are more familiar	3
They have more understanding of the topic of "social networking" than the "shared bikes".	2
The speech structure is helpful for the speech-making.	4
Providing glossaries in strategic planning is helpful.	7
Brainstorming in strategic planning is helpful.	6
Model speech in strategic planning is helpful.	5
Rehearsal is beneficial for speech-making.	6
The familiar topic can enhance spoken language production.	6
The reasons why the speech structures are helpful	3
The speech structure is related with the familiarity of the topic.	3
The reasons why rehearsal is good for language production	6
To compare the task conditions in Phase 1, 2, 3, the Phase 3 task conditions are the best for speech production.	5
The reasons why Phase 3 task conditions are the best	5
To compare the spoken language production in Phase 1, 2, 3, Phase 3 is the best.	4
The reason why Phase 3 spoken language production is the best	6

Table 5.16: The Most Common Nodes of the Initial Coding

An extract from focus group interview on the topic familiarity is as follows.

> Teacher: Last week, we have shared our views on the impact of shared-bicycles on our lives. Also, we recorded a speech on the travel experience in Shenzhen. I would like to know which topic you are more familiar with, and then could you explain the specific reasons for the two topics. Could each of us express our opinions in turn?
> Student 1: The shared-bicycles one.
> Teacher: Why?

Student 1: Because I haven't traveled in Shenzhen yet.
Teacher: Because everyone is busy studying, they haven't had time to go out, have they?
Student 1: It's mainly because I go home on National Day Holiday and don't bother to go out.
Teacher: Which topic: "the shared-bikes" or "travel experience in Shenzhen" can help you speak more fluent, accurate, and complex language? Why?
Student 1: Although they are not well spoken, it should be better for the "shared bicycles" topic.
Teacher: Speaking on familiar topics will be better, right?
Student 1: Yes.

When asked about the familiarity of topic, most students from the 6 participants considered the "shared bikes" topic in Phase 1 and the "social networking" topic in Phase 3 to be more familiar, because these topics were in their daily life and were easier to speak. As for the Phase 2 topic "describe a trip in Shenzhen", many of them thought this topic was demanding, as they had few travel experiences in Shenzhen at that time. As for the strategic planning, providing glossaries, model speeches, and brainstorming were considered useful in the speech preparation process. For the influence of task characteristics and task conditions on spoken language production, 6 quotes can be found, which are related to the familiar topic promoting speech production. In terms of the speech structure, there are 4 quotes from the students to support that it was helpful for speech-making. In the meantime, rehearsal was good for spoken language production. Regarding the task conditions of Phase 1, 2, and 3, it was supported from the students that the Phase 3 task conditions, with strategic planning and rehearsal, were better than the other two phases. Consequently, Phase 3 spoken language production was perceived as the best among the three research phases.

An extract from the focus group interview on strategic planning is as follows.

Teacher: When we talk about the "shared-bicycles" topic, we have brainstormed the content. We have also provided you with example speeches and useful expressions and vocabularies. Will these preparations help you in your speech production? Please tell me, the first is brainstorming, the second is providing a model speech, and the third is providing useful vocabularies. Which one helps you the best? Please explain with specific reasons.

Student 1: For me, it should be better to have a model speech, because it is possible that I can't find the way of thinking, or I don't know what to speak and the words to choose. Some words can also be found in the model.
Teacher: Are brainstorming and vocabularies useful or useless?
Student 2: I think they are both useful.
Student 3: I think those three are very useful. But if the model speech is released, I will not have my own ideas, and I may follow the speech.

Secondary Coding (Relational Login)

The main task of secondary coding (also known as relational login or axis login) is to find and establish various connections between conceptual categories, expressing the organic relationship between various parts of the data. These relationships can be causality, time sequence, semantic, similarity, difference, etc. (Chen, 1999). The secondary coding in this study is shown in Table 5.17. A sample of the focus group coding can be seen in Appendix 9.

The initial coding, which is related to the "shared bikes" topic, "describe a trip in Shenzhen", and "social networking", is categorized into the speech topics. Then, the "speech structure", "strategic planning", and "rehearsal" are labeled as the secondary coding. For the familiarity of topic, rehearsal, speech structure, and Phases 1, 2, 3, these four categories and their relationships with spoken language production are included in the secondary coding. The last two initial codes are generated as the "Phase 1, 2, 3 spoken language production".

An extract from the focus group interview on the 3 weeks of speech production is as follows. A sample of the focus group transcript can be found in Appendix 8.

Teacher: Among the 3 weeks of speech production with varied topics and task structure and plannings, which week do you have the best speech production?
Student 1: I think it's the third week. Brainstorming can make me think deeper. I think it's still necessary to rehearse. Of course, time is very important. I think the topic is more familiar.
Student 2: I think it must be the third week when I am fully prepared. For example, brainstorming can make us think more on the topic. Vocabulary can be equivalent to giving you an idea, thinking about how to say, and rehearsal can make us more familiar.
Student 3: I think it's the third week. I'm more prepared. On the other hand, this topic is

the most familiar among the three. So, in general, it is the third week.

Student 4: I think the third week is better. First, we have the brainstorming. In this way, we can have some general ideas about the main points and what words can be used in speeches and vocabularies, and the contents. Then, when we have time to organize the ideas, we can express more fluently. Because after all, after the rehearsal, we can speak more fluently.

Student 5: The third week should certainly be the most helpful. Brainstorming is to give ideas and improve the details of my speech. Rehearsing the speech can improve the whole speech process. So, I think the third week can be the most useful.

The Third Level Coding (Core Login)

The third level coding (also known as the core login or selective login) refers to selecting a "core category" after the systematic analysis of all discovered concept categories, where the analysis is constantly focused on the code numbers related to the core category (Chen, 1999). The core genera must repeatedly prove to be dominant in comparison with other genera, and can include the most research results in a relatively broad, theoretical scope. Like a fishing net, the core genera can lift all other genera as a whole and play the role of "outlining" (Chen, 1999). A set of scientific terms is used to establish the relationship between visible genera. With reference to terms, Strauss and Corbin (1998) have aggregated the expressions of the research objects and made it an integral part of the structural framework. In such a structural framework, they summarized these contents: 1. Conditions mean the environment or situation forming the structure of the studied subjects. 2. Actions indicate the routine or the strategic response of the research subjects to the theme, event, or problem. 3. Outcomes are the research results (Strauss & Corbin, 1998). Therefore, in this research, the speech topic, structure, strategic planning, and rehearsal can be regarded as the teaching conditions. The interventions of the research involving the different topics, structure, rehearsal, and task conditions are seen as the teaching actions. Consequently, the spoken language production is considered as the teaching outcomes (Table 5.17). To view this process in a more systematic way, the following figure (Figure 5.1) can represent the whole research.

The Third Level Coding	Secondary Coding		Initial Coding	Nodes
Teaching conditions	The speech topics	The shared bikes topic	The shared bikes topic is better.	11
			The understanding of the shared bikes topic is more.	3
			They have experience of writing the topic of shared bikes.	2
		The topic "describe a trip in Shenzhen"	They don't have much experience of traveling in Shenzhen.	7
			The topic "describe a trip in Shenzhen" is hard to express.	3
		The social networking topic	The topic social networking is more familiar.	7
			The reasons of why social networking topic are more familiar	3
			They have more understanding of the topic of social networking than the shared bikes.	2
	Speech structure		The speech structure is helpful for the speech-making.	4
	Strategic planning		Providing glossaries in strategic planning is helpful.	7
			Brainstorming in strategic planning is helpful.	6
			Model speech in strategic planning is helpful.	5
	Rehearsal		Rehearsal is beneficial for speech-making.	6
Teaching actions	Familiarity of topic and spoken language production		The familiar topic can enhance spoken language production.	6
	Speech structure and spoken language production		The reasons why the speech structures are helpful	3
			The speech structure is related with the familiarity of the topic.	3
	Rehearsal and spoken language production		The reasons why rehearsal is good for language production	6
	Phase 1, 2, 3 task conditions and spoken language production		To compare the task conditions in Phase 1,2, 3, the Phase 3 task conditions are the best for speech production.	5
			The reasons why Phase 3 task conditions are the best	5
Teaching outcomes	Phase 1, 2, 3 spoken language production		To compare the spoken language production, Phase 3 is the best.	4
			The reason why Phase 3 spoken language production is the best	6

Table 5.17: The Three Levels Coding

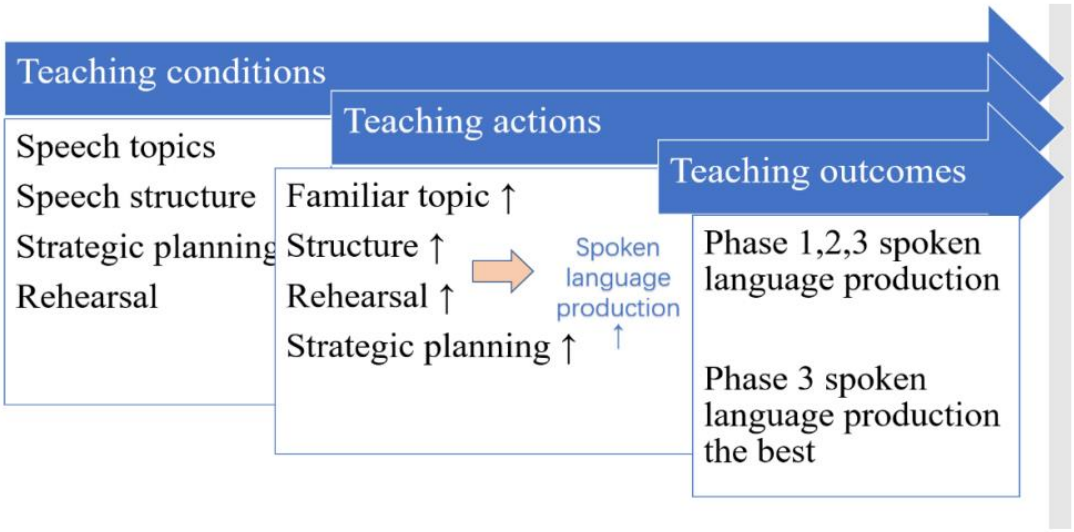


Figure 5.1: Structure Diagram of the Logical Relationship between the Main Categories and Genera

Theoretical Saturation Test

In this study, 1/4 of the interview records were reserved to test the theoretical saturation. The specific method is as follows: open login, relational login and core login are conducted again for the reserved 1/4 interview records. Then, whether new concepts would appear and whether the "genera" and "main genera" in the model would be saturated were observed.

Through inspection, it is found that the three main genera "teaching conditions", "teaching actions" and "teaching outcomes" have been saturated. No new important genera and relationships have been found, and no new constituent elements have been discovered in the initial coding and secondary coding, that is, the series of main genera have also become saturated. Therefore, the interview data of the influence of task characteristics and task conditions on students' spoken language production is saturated, in theory.

5.3 DATA TRIANGULATION

To answer the research question in this study, 3 types of data have been analyzed in this chapter. Below is a summary of the data of the 3 categories.

Quantitative Analysis of the Students' Recordings

Task characteristics	Familiarity of information ↑; Structure ↑	Lexical sophistication ↑
	Familiarity of information ↑; Structure ↑	Lexical density ↑
	Structure ↑	Syntactic complexity of clause per AS unit ↑
Task condition	Rehearsal ↑	Fluency of speech rate ↑

Table 5.15 Summary of Findings in Phase 1, 2 and 3 Research

Online Questionnaire Findings

In the general participant's viewpoints, the task characteristics: familiarity of topic and structure, and the task conditions: strategic planning and rehearsal, are beneficial for students' spoken language production in CAF.

Focus Group Interviews

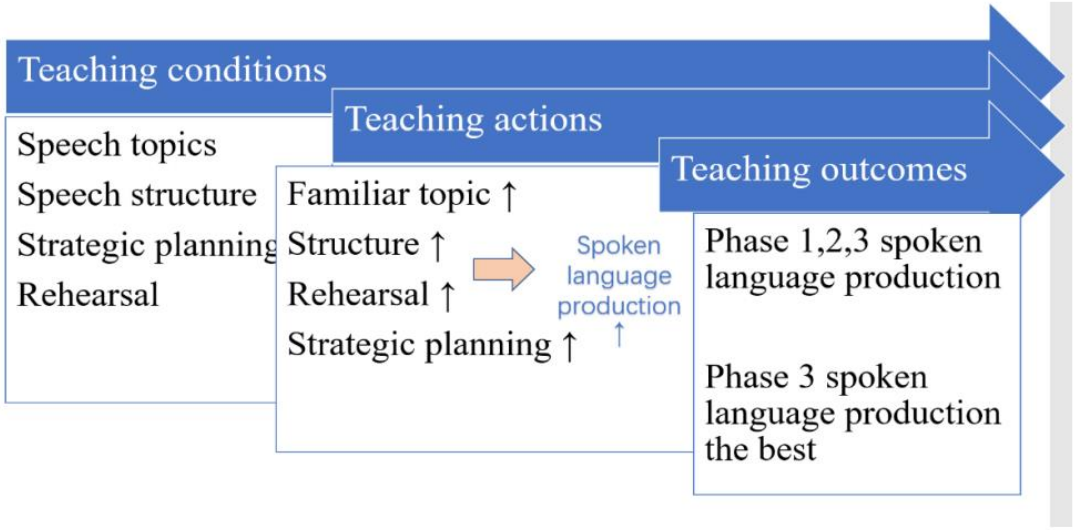


Figure 5.1: Structure Diagram of the Logical Relationship between the Main Categories and Genera

From the quantitative analysis, first, the familiarity of information and structure can benefit the lexical sophistication and density in spoken language production. Second, the task structure can promote the syntactic complexity of clause per AS unit. The task condition of rehearsal can increase fluency of speech rate.

Regarding the qualitative analysis, both the questionnaires and focus group interviews results demonstrate that task characteristics of familiarity of topic and task structure, as well as task conditions of strategic planning and rehearsal, are beneficial to the students' spoken language production. It is accepted by more than half of the online questionnaires' respondents and the focus group participants that with familiar topic, structure, strategic planning, and rehearsal in Phase 3, students could produce better speeches in CAF.

To summarize data triangulation, the interventions of task characteristics and task conditions are regarded as effective for spoken language production, in students' viewpoints. However, in the statistical analysis of quantitative data, familiarity of topic, structure, and rehearsal are partially related to the promotion of spoken language production. Next, the similarities and differences of findings between this study and previous research will be elaborated.

Similarities with Previous Findings

In line with previous researchers' findings, tasks with concrete or familiar information can facilitate accuracy and fluency. The structured tasks can increase accuracy in language production (Skehan, 2011). Rehearsal has no effect on CAF, but it is suggested that rehearsal is beneficial on the successive language performance of the same task (Ellis, 2013). In addition, strategic planning has shown positive effects on fluency and complexity.

The qualitative findings in this study show that the interventions of task characteristics and task conditions are effective for spoken language production in the qualitative analysis of students' viewpoints, supporting the views from Skehan (2011) and Ellis (2013) that familiarity of information enhances accuracy and fluency, while degree of structure promotes accuracy and strategic planning helps complexity and fluency. Moreover, from Ellis's (2005) analysis, it is suggested that rehearsal is beneficial to the successive language performance of the same task, which is supported from the qualitative analysis of students' viewpoints and partly proved by the quantitative analysis of rehearsal, enhancing the fluency of the speech rate. The reasons for this agreement probably lie in the subjective concepts of the students, considering the effectiveness of the variables in task characteristics and task conditions on CAF.

Differences with Previous Findings

However, in the statistical analysis of quantitative findings, familiarity of topic and structure seem to be beneficial to the lexical and syntactic complexity. Rehearsal can foster the fluency of speech rate. These are not in alignment with the claims of Skehan (2011) and Ellis (2013)'s previous research. To explain the contradictions, the variables in this study are not tested alone in the three phases of research and one variable in task characteristics and task conditions could influence the other in terms of spoken language production in CAF.

5.4 CONCLUSION

This chapter has analyzed the quantitative data of students' recordings by the test of normality, paired t-test, and nonparametric tests for the influence of task condition of rehearsal in the students' speech production. For the task characteristics, the repeated measures of ANOVA have been adopted to examine the familiarity of topic and structure and their influence on spoken language production. Then, the qualitative

data of questionnaires have been analyzed for the answers to the research question. For the focus group interviews, the grounded theory approach has been applied to interpret the results of the research question. The qualitative findings show that the interventions of task characteristics and task conditions are regarded as effective for spoken language production, in students' viewpoints. However, in the statistical analysis of quantitative data, familiarity of topic, structure, and rehearsal are partially related to the promotion of spoken language production. Based on the findings in this chapter, further discussion of the research question will be presented in the next chapter.

Chapter 6 Conclusion

In this chapter, the themes of this study, research question, and research findings are summarized. These shed light upon the application of the speech tasks in a College English course, English Listening and Speaking, including the influence of task characteristics and task conditions on students' spoken language production in a university of science and technology, in China. The contributions to knowledge and the perceived implications for teachers and classroom practice are reported. Meanwhile, the research limitations are discussed and areas for further research are suggested. Finally, the concluding remarks related to the personal experience obtained from this research process are presented at the end of this chapter.

6.1 THEMES OF THE STUDY

This research has focused on the three major themes: 1. tasks; 2. spoken language production; 3. task and spoken language production, including the relationship between tasks characteristics, tasks conditions and spoken language production.

Theme 1: Tasks

As mentioned in Chapter 1 Introduction, the definition of a task is demonstrated: a task is goal-oriented (Candlin, 1987), with a number of steps, which follows a series of cognitive and communicative procedures, and has a defined outcome (Ellis, 2013). Additionally, a task is sequential and can be subject to pedagogical interventions (Prabu, 1987). As for the task types, the strong types of tasks with purposeful, authentic communication, and open-ended outcomes were chosen, as this study sought meaningful and real-life communication from the learners.

Theme 2: Spoken Language Production

For L1 production, Levelt (1989) has proposed the information processing model in three hierarchically modular stages: conceptualization, formulation, and articulation.

The conceptualization stage is to develop and organize the ideas to a communicative goal. Then, in the formulation stage, a phonetic plan is made for the content of speaking. In the end, articulation is created when the phonetic plan is transformed into the actual speech (Ellis, 2013).

However, L2 production in this study is complicated compared to L1 production. The concepts of L2 production, the understandings of learners' SLA in the Input and Output Hypotheses have been discussed in Chapter 3. Based on the Output Hypothesis, the triad constructs of Complexity, Accuracy, and Fluency (CAF) are identified by Skehan (1999), which can serve as the measures of second language production.

CAF Constructs

The three perspectives of L2 production have been distinguished by Skehan (1999). CAF have been viewed as the principal research variables of language production in L2 research (Ellis, 2009), which are defined in Table 6.1.

Complexity	The capacity to use more advanced language, with the possibility that such language may not be controlled effectively. This may also involve a greater willingness to take risks, and use fewer controlled language subsystems.
Accuracy	The ability to avoid errors in performance, possibly reflecting higher levels of control in language, as well as a conservative orientation.
Fluency	The capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems.

Table 6.1: Definitions of Complexity, Accuracy and Fluency (Skehan & Foster, 1999)

Nevertheless, unlike L1 production, L2 speech production can be more demanding in the cognitive information process along with insufficient and limited mental lexicon, and the storage of "considerable information about each lemma and information", to support the natural and immediate L1 speech production (Skehan, 2011: 253).

Based on the limited resources of attention and working memory in L2, it is natural

for the learners to focus on meaning instead of form in a communicative context (VanPatten, 2007). The possibility could be that the form is lost at the expense of advanced language to achieve the primary goals of fluency and meaningful expressions (Skehan, 2014). Focusing on the cognitive perspectives, Skehan (2014) has proposed the Limited Attentional Capacity Model and emphasized the values of attention and working memory and proposed the framework which indicates that learners' cognitive capacity of attentional working memory resources is limited to achieve the speech production in CAF. These assumptions were adopted as the theoretical framework in this study.

Theme 3: Task and Spoken Language Production

Task Characteristics and Spoken Language Production

As for the relationship between task characteristics and spoken language production, Skehan (2014) has identified the tasks features including familiarity of information (concrete-abstract; familiar-unfamiliar material) and degree of structure (structured-unstructured) in the task design process. For familiarity of information, the tasks "vary as to whether they require information that is familiar to the participants as part of their personal experience" (Skehan, 2011: 235). For the degree of structure, some tasks have a clear and over-arching structure while some do not (Skehan, 2011). The task characteristics and their relationships with spoken language production can be summarized as follows (Table 6.2).

Task characteristics	Familiarity of information	Degree of structure
Spoken Language production	Familiarity↑ A ↑ F ↑ C-	Structure↑ A ↑

Table 6.2: The Relationship between Task Characteristics and Spoken Language Production (Skehan, 2011)

In this study, the varied familiarity of information and degree of structure in three phases of the speech tasks are scrutinized for their relationship with spoken language

production in the research context.

Task Conditions and Spoken Language Production

In this study, the speech-making task has been selected. In the pre-task planning phase, the task conditions of implementation have received focus, as teachers can apply pedagogical interventions in the task preparation process.

Pre-task Planning

The pre-task planning phrase, which is the planning before learners perform a task, is identified by Ellis (2005) in two aspects: rehearsal and strategic planning (Table 6.3).

Pre-task planning is planning that is done before learners perform a task.	Rehearsal	Planning takes the form of an opportunity to perform the complete task once before performing it a second time.
	Strategic planning	Planning includes the contents to be expressed and the language to be used but without an opportunity to rehearse the complete task.

Table 6.3: Pre-task Planning and Types of Pre-task Planning (Ellis, 2009: 474)

Ellis (2013) has demonstrated the relationship of rehearsal and strategic planning to language production in Table 6.4. In this study, the relationship of rehearsal and strategic planning in the three phases of the speech tasks have been investigated with spoken language production.

Language performance	Rehearsal	Strategic planning
Fluency	No effect	Positive effect
Accuracy	No effect	Effects sometimes evident
Complexity	No effect	Positive effect

Table 6.4: The Effects of Rehearsal and Strategic Planning on Language Performance (Ellis, 2013: 133)

Research Question

How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?

As outlined and presented in Chapter 4, this research study has answered the above research question with the objectives set to drive the research undertaken by the quantitative use of the three constructs: CAF measures in students' speech recordings, along with the qualitative analysis of online questionnaires and focus group interviews. These constituted the basis of a research design rationale, with the foundation of a pragmatism paradigm and mixed methods research substantiated in the same chapter. For the sake of clarity, these research findings are presented.

As mentioned in Chapter 5, this thesis provides findings from the quantitative analysis that the familiarity of information and structure can be partly beneficial to the lexical complexity in spoken language production. The task structure can promote syntactic complexity. While the task condition of rehearsal can increase fluency of speech rate. The summary of quantitative findings is as follows (Table 6.5).

Task characteristics	Familiarity of information ↑; Structure ↑	Lexical sophistication ↑
	Familiarity of information ↑; Structure ↑	Lexical density ↑
	Structure ↑	Syntactic complexity of clause per AS unit ↑
Task condition	Rehearsal ↑	Fluency of speech rate ↑

Table 6.5: Summary of Quantitative Findings

Regarding the qualitative analysis, both the questionnaires and focus group interviews results demonstrate that the task characteristics of familiarity of topic and task structure, as well as the task conditions of strategic planning and rehearsal can promote students' spoken language production. It was reported by more than half of the online questionnaires' respondents and the focus group participants that with familiar topic, structure, strategic planning, and rehearsal, students could produce

better speech in CAF.

6.2 CONTRIBUTIONS TO KNOWLEDGE

Contributions to Previous Studies

This study has contributed to the significant pedagogy topic on the theoretical framework of task characteristics and task conditions and spoken language production in teaching English speaking in a university of science and technology.

From the previous studies, Skehan (2011) and Ellis (2013) have suggested that familiarity of information enhances accuracy and fluency, while degree of structure promotes accuracy and strategic planning helps complexity and fluency. Moreover, from Ellis's (2005) analysis, it is suggested that rehearsal is beneficial to the successive language performance of the same task. In a more recent study, Ellis, Li, and Yan (2019) have found that pre-task instruction can lead to more fluent but not more accurate use of language and that it had detrimental global effects on the complexity, accuracy, and fluency of the learners' production. Another recent study by Ellis (2022) indicates that pre-task planning does not lend unconditional support on language learners in designing tasks.

In this study, the findings show that the interventions of task characteristics and task conditions are effective for spoken language production in the qualitative analysis of students' viewpoints, supporting the views from Skehan (2011), Ellis (2013), Ellis, Li, and Yan (2019) that familiarity of information enhances accuracy and fluency, degree of structure promotes accuracy and strategic planning helps complexity and fluency, pre-task planning of task conditions improve fluency (Table 6.5). Moreover, from Ellis's (2005) analysis, it is suggested that rehearsal is beneficial to the successive language performance of the same task, which is supported from the qualitative analysis of students' viewpoints and partly proved by the quantitative analysis of rehearsal, enhancing the fluency of the speech rate.

However, in the statistical analysis of quantitative findings, familiarity of topic and structure are suggestive to be beneficial to the lexical and syntactic complexity. Rehearsal can foster the fluency of speech rate. These are not in total alignment with the claims of Skehan (2011) and Ellis' (2005, 2013, 2022) previous research (Table 6.6). To explain the contradictions, the variables of task characteristics and task conditions in this study are not tested alone in the three phases of research and one variable could influence the other in terms of spoken language production in CAF. Still, this study can be regarded as a reference for further research in the future.

<table>
<tr><td colspan="5">Theoretical Framework (Skehan, 2011; Ellis, 2005/2013/2022; Ellis, Li & Yan, 2019)</td></tr>
<tr><td rowspan="2">Task characteristics</td><td colspan="2">Familiarity of information ↑</td><td colspan="2">Degree of structure↑</td></tr>
<tr><td colspan="2">A ↑ F ↑ C-</td><td colspan="2">A ↑</td></tr>
<tr><td rowspan="4">Task conditions</td><td colspan="2">Rehearsal ↑</td><td colspan="2">Strategic planning↑</td></tr>
<tr><td>C- A- F-</td><td>F↑ A-</td><td colspan="2" rowspan="2">C↑ F↑</td></tr>
<tr><td>the successive language performance of the same task ↑</td></tr>
<tr><td colspan="3">Pre-task plannings do not lend unconditional support.</td></tr>
<tr><td colspan="5">Research Findings of this Study</td></tr>
<tr><td colspan="5">Quantitative Findings</td></tr>
<tr><td>Familiarity of information ↑; Structure ↑</td><td colspan="2">Familiarity of information ↑; Structure ↑</td><td>Structure ↑</td><td>Rehearsal ↑</td></tr>
<tr><td>Lexical Sophistication ↑</td><td colspan="2">Lexical density ↑</td><td>Syntactic complexity of clause per AS unit ↑</td><td>Fluency of speech rate ↑</td></tr>
<tr><td colspan="5">Qualitative Findings</td></tr>
<tr><td>Familiarity of information ↑</td><td colspan="2">Degree of structure ↑</td><td>Rehearsal ↑</td><td>Strategic planning ↑</td></tr>
<tr><td>C↑ A ↑ F↑</td><td colspan="2">C↑ A ↑ F↑</td><td>C↑ A ↑ F↑</td><td>C↑ A ↑ F↑</td></tr>
</table>

Table 6.6: A Comparison of the Theorical Framework and the Research Findings

Originality of this Study

This research study has focused on speaking and set in the specific course of English Listening and Speaking under the CE curriculum. As mentioned in Chapter 2, speaking is a weak link in ELT in China and it is a high-stakes proficiency skill for

language learners. Furthermore, there exists a tension between speaking skill and career choices for language learners, especially for the engineering students from different majors in this study. The findings in this study can provide a reference for the context of the target students in teaching speaking for classroom practitioners. Meanwhile, the pedagogical implications for teaching speaking in the context of this study can be applied to many different contexts in Chinese ELT.

Second, this study has supplemented the current theoretical framework of task characteristics and task conditions on students' spoken language production with the quantitative findings. The research question, "How do different task characteristics and task conditions impact students' spoken language production (CAF) in English Listening and Speaking?", is answered by both the quantitative and qualitative findings. From the quantitative analysis that the familiarity of information and structure can be partly beneficial to the lexical complexity in spoken language production. The task structure can promote syntactic complexity, while the task condition of rehearsal can increase fluency of speech rate. In this study, the quantitative findings serve as the central findings and supplement to the current theoretical framework (see Table 6.6). Considering the qualitative analysis, the findings are subjective from the students, which can be regarded as arising from the background and contextuality of this study.

Third, the robustness of the methodology stems from the adoption of the explanatory model of MMR in data triangulation from the quantitative and qualitative findings for the answers to the relationship of task characteristics and task conditions on spoken language production. The quantitative analysis has involved SPSS statistical analysis, while the qualitative analysis of the focus group interviewees has applied the grounded theory, which provides valuable, reliable, and valid findings by means of research methods. The significance in the research findings has been demonstrated by

the rigorous methods in SPSS analysis and the grounded theory, which can be useful in the current literature and practice.

Fourth, the research study can be generalized to other contexts worldwide. For instance, the target students are weak in speaking and shy in their characteristics, which is not exclusive to Chinese students. Furthermore, the CE curriculum of English Listening and Speaking is limited in teaching time and faced with many challenges with the application of TBLT. These limitations and challenges are not unique to China.

Last but not the least, the results of this study can be valuable for future practitioners and researchers who would like to focus on TBLT in speaking classes either in their pedagogical practice or in their further research.

6.3 IMPLICATIONS FOR PRACTICE

First, this study has multiple implications for CE teachers:

- It provides an opportunity for teachers of CE speaking classes to reflect on their own teaching practice with regards to the quantitative recordings of students' CAF measures, qualitative questionnaires, and focus group interviews. From the quantitative analysis, it appears that the familiarity of information and structure can be partly beneficial to the lexical complexity in spoken language production. The task structure can promote syntactic complexity, while the task condition of rehearsal can increase fluency of speech rate. For language teachers, more familiar topics can be chosen to facilitate students' language production in lexical complexity. Providing speech structure can be helpful for both lexical and syntactic complexity. For promoting fluency, rehearsal can be arranged in pre-task planning. As it is indicated in the qualitative findings, the more familiar topics, and the more structured tasks in task characteristics, along with strategic planning and rehearsal in task conditions could promote spoken language

production. Based on these, teachers can select a familiar topic, with the scaffolding of speech structure, when designing a speech task, and implement strategic planning and rehearsal to promote students' spoken language production.

- It encourages teachers to reflect on their own context of teaching by designing and implementing more tailored tasks for their target students. The teaching of speaking for engineering students from the university of science and technology can be different in task design and implementation from students of other contexts, for example, the English majors in a comprehensive university.

- It creates awareness that task design and implementation are influenced by various factors such as familiarity of information, degree of structure, rehearsal, and strategic planning. This awareness may prompt teachers to provide more familiar and more structured tasks along with strategic planning and rehearsal in the task characteristics and task conditions and provide fewer demanding tasks for students with limited cognitive attention and working memory.

Second, this study may have implications for universities of science and technology in the teaching of General English courses in CE by offering awareness and understanding that:

- CE General English courses should be taught at different levels. In different levels of teaching, teachers can design the speaking tasks according to students' English levels and speaking proficiency, and implement them in the classroom with suitable task characteristics in design and task conditions in implementation.

Third, this study may have implications for government policy makers by encouraging curriculum planners to consider a more comprehensive CE curriculum

that would accommodate the learning needs of the diverse students in colleges and universities in China. Taking the university of science and technology as an example, it is necessary to enhance international awareness of students by broadening their horizons, and improving intercultural communication and competitiveness in response to the "Excellent Engineer Education and Training Program" raised by Ministry of Education in 2010 (Sheng, 2015). Such training should be inseparable from the development of CE curriculum for these students. Therefore, in addition to a certain number of General English courses for English skills of listening, speaking, reading, and writing, corresponding ESP courses and Intercultural Communication courses related to students' professional engineering needs should be devised in this context to ensure the instrumental and humanistic nature of the CE curriculum (Sheng, 2015).

Finally, this study can be a pedagogical reference for language teachers in international contexts. As it is mentioned in the originality of this study, the target students are weak in speaking, shy in their characteristics and passive in language classroom, which is not exclusive to English learners in China. Similar cases can be identified with Chinese learners in Korea and Thailand (Jin, 2011; Yan, Guo & Chen, 2020). Teachers in these countries with the same target students can apply TBLT in language teaching and the empirical experiences in this study can be applicable to enhance students' spoken language production. On the other hand, the CE curriculum of English Listening and Speaking is limited in teaching time and faced with many challenges with the application of TBLT, which are problems that can also be found in Chinese learners in Korea and Thailand (Jin, 2011; Yan, Guo & Chen, 2020).

6.4 LIMITATIONS OF THE STUDY

This study has focused on two elements of task characteristics: familiarity of information and task structure with their influences on spoken language production, which is not the whole picture of task characteristics. To take other perspectives into account, dialogic tasks and complex outcomes of tasks can be studied further in task

characteristics of spoken language production (Skehan, 2011). Respecting the task conditions, this study concentrates merely on the pre-task planning without looking further into within-task planning (Ellis, 2005), which can be limited in the task implementation process.

In addition, this study concerns only the first-year university students in the class of English Listening and Speaking in a university of science and technology in China. On one hand, the sample group for quantitative analysis contains 30 students, which was restricted in the quantity. On the other hand, as the context was established in a university of science and technology in China for a General English course in CE curriculum, the findings could be confined to this specific context, instead of a wider and more general research context.

Furthermore, the quantitative data analysis process could be constrained by the paired t-tests, nonparametric tests, and repeated measures of ANOVA analysis of SPSS in the speech recordings. Alternative methods of statistical analysis could possibly yield other quantitative results. Regarding the focus group interviews, it would be merely the opinions of 6 students. Their opinions for the research question could be limited by the time and space of the data collection for the focus group participants.

6.5 FURTHER STUDY EMERGING FROM THE THESIS

Task Characteristics and Task Conditions

This study has explored the two task characteristics: familiarity of information and degree of structure, along with the varied task conditions: rehearsal and strategic planning on students' spoken language production in the General English course of English Listening and Speaking. Further research can include the varied familiarity of information, degree of structure, strategic planning, and rehearsal, each alone in the research, so that researchers can see a clearer picture of each variable and their influence on students' spoken language production.

As mentioned above, in limitations, the dialogical tasks and complex outcomes of tasks could be studied further in task characteristics. The different types of within-task planning, including pressured and unpressured tasks (Ellis, 2005), could be emphasized to investigate the task conditions in students' spoken language production. Apart from speaking, future research could focus on reading, writing, and listening in the General English courses to apply the varied task characteristics and task conditions on students' language production.

Limited Attentional Capacity Model

The study is based on the framework of the Limited Attentional Capacity Model by Skehan (2014). It is believed that the information processing ability for students is limited, so it is necessary to optimize the distribution of their attention (Skehan, 1998). In other words, if a task needs to invest more attentional resources in meaning, less attentional resources will be paid in the language form. Language learners will give priority to allocating attentional resources to the meaning of information rather than language form (Yi, 2014). In this way, there is a competition for attentional resources in form and meaning. Paying attention to meaning will promote language fluency, while paying attention to form will promote language accuracy and complexity. Therefore, it comes with the "trade-off" effect in language production (Yi, 2014).

In contrast with Skehan's (2014) Limited Attentional Capacity Model, Robinson (2007) has proposed the Cognition Hypothesis that learners can mobilize multiple resources at the same time when completing complex tasks, so there may not be competition for attentional resources among CAF in language production. However, in recent years, more empirical studies, especially psychological studies have confirmed that Skehan's (2009) cognitive model of language production is more explanatory (Ellis, 2009), and Levelt's Speech production model also applies to the limited attentional resources hypothesis (Yi, 2014).

Nevertheless, further research could be designed in classroom to test the effectiveness between Skehan' s Limited Attentional Capacity Model and Robinson's Cognition Hypothesis to generate answers in task design and task implementation on the spoken language production.

Literature Course

This study began with the exploration of task characteristics and task conditions on students' spoken language production in a British and American Literature course. With the spread of COVID-19 pandemic, the original design could not be realized at that time. Future study, however, can be continued to explore the task design and task implementation in the CE curriculum of British and American Literature.

6.6 CONCLUDING REMARKS

This study contributes to the theoretical knowledge of the relationships of task characteristics, task conditions and spoken language production. The research findings have provided reference and given rise to possible research in the future.

During the research process, I have had the opportunity to understand and learn much about the issues of task characteristics, task conditions and spoken language production. In addition, I have acquired valuable skills as a pedagogical researcher in designing the speech tasks with varied task characteristics and task conditions, conducting them in my own classes, gathering and analyzing both the quantitative and qualitative data, and presenting the findings, which will be invaluable for my future research foundations and professional development. Nevertheless, the research limitations and difficulties outlined in this chapter illustrate the necessity of further research in respect to the fuller and more comprehensive answers of the research issues in task design and task implementation.

The efforts in this research, with the enhancement of knowledge and pedagogical

practice, has offered classroom recommendations and included some practical directions. Future study should be conducted, with this work serving as a springboard and valuable reference. Even though this research comes to an end, the search for truth will never end. The end of this study can mark the beginning of a new chapter in my journey as a pedagogical researcher.

References

Adamson, B. (2004a). *China's English: A History of English in Chinese Education*. Hong Kong University Press.

Adamson, B. (2004b). Fashions in Language Teaching Methodology. In A. Davies & C. Elder (Eds.), *The Handbook of Applied Linguistics*. Blackwell Publishing Ltd.

Babbie, E. (2017). *The Practice of Social Research (11th Edition)*. Beijing: Tsinghua University Press.

Bao, Y. (2019). Strategies for Improving College Students' Speech Ability. *Journal of Taiyuan Urban Vocational College*, (012), 150-152.

Biber, D., Gray, B., & Staples, S. (2016). Predicting Patterns of Grammatical Complexity across Language Exam Task Types and Proficiency Levels. *Applied Linguistics*, (37)5: 639-668.

Biddle, C. & Schafft, K. A. (2015). Axiology and Anomaly in the Practice of Mixed Methods Work. *Journal of Mixed Methods Research*, 9(4), 320-334.

Breen, M. (1989). The Evaluation Cycle for Language Learning Tasks. In R. K. Johnson (Ed.), *The Second Language Curriculum*. Cambridge: Cambridge University Press.

Brumfit, C. (1984). *Communicative Methodology in Language Teaching*. Cambridge: Cambridge University Press.

Bryman, A. (2006). Paradigm Peace and the Implications for Quality. *International Journal of Social Research Methodology, 9*(2), 111-126.

Bryman, A. (2008). *Social Research Methods (3rd Edition)*. Oxford: Oxford University Press.

Bui, G. (2014). Chapter 3. Task Readiness: Theoretical Framework and Empirical Evidence from Topic Familiarity, Strategic Planning, and Proficiency Levels. In P. Skehan. (Ed.). *Processing Perspectives on Task Performance*. Amsterdam: John Benjamins Publishing Company.

Bulté, B. & Housen, A. (2012). Defining and Operationalizing L2 Complexity. In A.

Housen, F. Kuiken& I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity Accuracy and Fluency in SLA*. Amsterdam: John Benjamins.

Burrell, G. & Morgan, G. (1979). *Sociological Paradigms and Organizational Analysis*. London: Heinemann Educational.

Bygate, M. (1987). *Speaking. Oxford: Oxford University Press.*

Bygate, M. (1988). Units of Oral Expression and Language Learning in Small Group Interaction. *Applied Linguistics*, (1), 59-82.

Bygate, M. (1996). Effects of Task Repetition: Appraising the Developing Language of Learners. In J. Willis and D. Willis (Eds), *Challenge and Change in Language Teaching*. Oxford: Heinemann.

Bygate, M. (2001). Effects of Task Repetition on the Structure and Control of Oral Language. In M. Bygate, P. Skehan and M. Swain (Eds), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing.* Harlow: Longman.

Bygate, M.P., Skehan, P., & Swain, M. (2001). *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing.* Harlow: Longman.

Cai, J. G. (2015). A Literature Review of Four Positioning Debates in the History of College English Teaching in China and Its Implications. *China University Teaching,* (10), 36-42.

Cai, L.Y. (2019). Research on the Application of Cooperative Teaching in College English Viewing, Listening and Speaking Classroom. *Science Mass (Science Education),* (02), 142

Candlin, C. N. (1987). Towards Task-based Language Learning. In C. N. Candlin & D. F. Murphy (Eds.), *Language Learning Tasks.* Englewood Cliffs, NJ: Prentice Hall International.

Chang, Y. J. & Luo, L. (2005). The Controversy of Research Paradigms and Its Enlightenment to Research. *Studies in Dialectics of Nature,* 21(7), 20-23.

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis.* London: Sage.

Chen, H.Y. (2008). Task-based Approach to Teaching: From Theory to Practice. *Foreign Language Education in China,* (2), 11-18.

Chen, X. M. (1999). Ideas and Methods on the Grounded Theory. *Educational Research and Experiment*, (004), 58-63.

Chen, X. M. (2015). The Application of Grounded Theory in Chinese Educational Research. *Peking University Education Review*, 13(01), 2-15.

Clark, H. & Clark, G. (1977). *Psychology and Language: An Introduction to Psycholinguistics*. New York: Harcourt, Brace, Jovanovich.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37-46.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education (7th Edition)*. London and New York: Routledge.

Cortazzi, M. & Jin, L. X. (1996). English Teaching in China. *Language Teaching*, *29*, 61-80.

Creswell, J.W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd Edition)*. Thousand Oaks, CA: Sage.

Creswell, J.W. (2014). *A Concise Introduction to Mixed Methods Research*. Thousand Oaks, CA: Sage.

Creswell, J.W. & Tashakkori, A. (2007). The New Era of Mixed Methods. *Journal of Mixed Methods Research*, 1(1), 3-7.

Creswell, J.W. & Plano-Clark, V.L. (2007/2011). *Designing and Conducting Mixed Methods Research (1st/2nd Edition)*. Thousand Oaks, CA: Sage.

Creswell, J.W., Plano-Clark, V.L., Gutman, M.L., & Hanson, W.E. (2003) Advanced Mixed Methods Research Designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks, CA: Sage. 209-240.

Crookes, G. (1986). Task Classification: A Cross-disciplinary Review. Technical Report No.4. Honolulu: Center for Second Language Classroom Research Social Science Research Institute. University of Hawaii at Manoa.

Crystal, D. (1997). *English as a Global Language.* Cambridge: Cambridge University Press.

Cui, G. Y. (2021). A Survey of Foreign Language Learning Anxiety among Chinese and Croatian College Students. *Foreign Language Education* (01), 3-17

Dai, Y. C. & Dai, W. D. (2010). An Analysis of the Acquisition Process from Input to Output and its Psychological Mechanism. *Foreign Language World*, (01), 23-30.

Daneman, M. & Carpenter, P. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior,* (19), 450-466.

Dell, G. S. (1986). A Spreading-activation Theory of Retrieval in Sentence Production. *Psychological Review, 93*(3), 283-321.

Deng, M. & Pan, J. F. (2002). On the Design of Mixed Methods in Educational Research. *Educational Research and Experiment*, (003), 56-61.

Denscombe, M. (2008). Communities of Practice: A Research Paradigm for the Mixed Methods Approach. *Journal of Mixed Methods Research, 2*(3), 270-283.

Denzin, N. K. (2010). Moments, Mixed Methods, and Paradigm Dialogs. *Qualitative Inquiry, 16*(6), 419-427.

Ding, H. & Dai, W. P. (2008). Chomsky's Theory of Language Innatism. *Social Sciences in Guangxi,* (07), 144-147.

Doyle, L., Brady, A. M., & Byrne, G. (2016). An Overview of Mixed Methods Research-Revisited. *Journal of Research in Nursing*, *21*(8), 623-635.

Ellis, R. (1987). Interlanguage Variability in Narrative Discourse: Style Shifting in the Use of the Past Tense. *Studies in Second Language Acquisition, 9*(1), 1-19.

Ellis, R. (2005). Planning and Task-based Research: Theory and Research. In R. Ellis (Ed.): *Planning and Task-Performance in a Second Language*. Philadelphia: John Benjamins.

Ellis, R. (2009). The Differential Effects of Three Types of Task Planning on the Fluency, Complexity, and Accuracy in L2 Oral Production. *Applied Linguistics, 30*(4), 474-509.

Ellis, R. (2013). *Task-based Language Learning and Teaching*. Shanghai: Shanghai

Foreign Language Education Press.

Ellis, R. (2022). Does Planning before Writing Help? Options for Pre-task Planning in the Teaching of Writing. *ELT Journal, 76*(1): 77-87.

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

Ellis, R., Li, S., & Yan, Z. (2019). The Effects of Pre-task Explicit Instruction on the Performance of a Focused Task. *System*, 80.

Fang, W. L. (2003). On Task-based Language Teaching. *Foreign Languages and Their Teaching,* (009), 17-20.

Feilzer, M. Y. (2010). Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm. *Journal of Mixed Methods Research, 4*(1), 6-16.

Feng, Y. F. & Tang, X. Y. (2004). The Application of Task-based Language Teaching Method in English Teaching. *Foreign Languages and Their Teaching*, (06), 35-38.

Foster, P. et al. (2000). Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics, 21*(3), 354-375.

Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition, 18*(3), 299-323.

Foster, P., & Skehan, P. (1999). The Influence of Source of Planning and Focus of Planning on Task-based Performance. *Language Teaching Research, 3*(3), 215-247.

Fromkin, V. A. (1973). *Speech Errors as Linguistic Evidence.* The Hague: Mouton.

Gage, N. (1989). The Paradigm Wars and their Aftermath. *Educational Researcher, 18*(7), 4-10.

Gil, J., & Adamson, B. (2011). The English Language in Mainland China: A Sociolinguistic Profile. In A. Feng. (Ed.), *English Language Education across Greater China* (pp. 23-45). Buffalo, NY: Multilingual Matters.

Greene, J. C. (1990). Technical Quality Versus User Responsiveness in Evaluation

Practice. *Evaluation & Program Planning*, *13*(3), 267-274.

Greene, J. C., Caracelli, V. J., & Graham, W. D. (1989). Toward a Conceptual Framework for Mixed-method Evaluation Designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255-274.

Gu, Q.Y., & Jin, X. F. (2021). Cognitive Load in Chinese EFL Learners' Output: The Role of Output Modality and Topic Familiarity. *Foreign Languages*, (3), 9

Guba, E. G. (1990). *The Paradigm Dialog.* Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1994). Competing Paradigms in Qualitative Research. In N. K. Denzin & Y.S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105-117). Thousand Oaks, CA: Sage.

Guo, Y. (2016). *The Change of English Education in China and its Influence on Social Development.* Guangming Daily Press.

Hall, E. (2010). *Beyond Culture*. Beijing: Beijing University Press.

Hammersley, M. (2012) Methodological Paradigms in Educational Research, British Educational Research Association On-line Resource. Available on-line at [http://www.bera.ac.uk/resources/methodological] Last accessed [February 12th, 2023]

Hanson, B. (2008). Wither Qualitative/quantitative? Grounds for Methodological Convergence. *Quality &Quantity,* 42, 97-111.

He, D.Z., Li, T & Tao, Y. C. (2017). College Oral English Teaching Strategies Based on Cultural Differences between China and the West. *Journal of Liaoning University of Technology (Social Science Edition), 19*(001), 114-116.

He, Y. & Shi, D. S. (2009). Discussion on the Research Method of "Grounded Theory" in Social Survey. *The World of Survey and Research,* (05),46-48.

Housen, A. & Kuiken, F. (2009). Complexity, Accuracy and Fluency in Second Language Acquisition. *Applied Linguistics,* (30): 461-473.

Hu, X. Q. (2007). The Development Trend of "Global English" in the Globalization Environment and its Enlightenment to English Teaching in China. *Foreign Language Teaching Abroad,* (2), 16-23.

Izumi, S. et al. (1999). Testing the Output Hypothesis: Effects of Output on Noticing and Second Language Acquisition. *Studies in Second Language Acquisition*, *21*(3), 421-452.

Jin, Y. & Yang, H. Z. (2018). Taking the Path of Language Testing with Chinese Characteristics: The Enlightenment of CET-4 and CET-6 for 30 Years. *Foreign Language World*, (002), 29-39.

Jin, Y. S (2011). Research on the Application of Task-based Teaching method in Korean High School Oral Chinese Class. (Master thesis, Shandong Normal University)

Johnson et al. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Johnson, R. B. & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33, 14-26.

Kachru, B. (1985). Standards, Codification and Sociolinguistic Realism: The English Language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.). *English in the World: Teaching and Learning the Language and Literatures.* Cambridge: Cambridge University Press.

Kerr, A. W., Hall, H. K. & Kozub, S. A. (2002). *Doing Statistics with SPSS.* Sage Publications.

Knight, P. (2001). The Development of EFL Methodology. In C. N. Candlin & N. Mercer (Eds.), *English Language Teaching in Its Social Context* (pp. 147–66). London: Routledge.

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

Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon.

Kuhn, T. S. (1962). *The Structure of Scientific Evolutions.* Chicago: The University of Chicago Press.

Lam, A. (2002). English in Education in China: Policy Changes and Learners' Experiences. *World Englishes, 21*(2), 245-256.

Lan, Y. C. (2009). The Application of Task-based Approach in College Oral English Teaching. *Journal of Lanzhou University (Social Science Edition)* (S1), 153-156.

Lee, J. (2000). *Tasks and Communicating in Language Classrooms*. Boston: McGraw-Hill.

Legutke, M. & Thomas, H. (1991). *Process and Experience in the Language Classroom*. Longman Group UK Limited.

Levelt, W. J. (1989). *Speaking: From Intention to Articulation*. Cambridge: Cambridge University Press.

Levelt, W. J. (2008). *Speaking: From Intention to Articulation*. Beijing: Foreign Language Teaching and Research Press.

Li, B., Xing, F.& Wang, H. (2019). A Review and Analysis of the Development and Research of College English Education since the Founding of PRC. *Journal of Hebei Normal University (Education Science Edition), 021*(003), 56-64.

Li, C. (2011). Implementation of ANOVA of Repeated Measurement Data in SPSS 16.0 Software. *Statistics and Decision, 000*(001), 34-36.

Li, H. (2002). The Cognitive Basis of Comprehensible Output Hypothesis. *Foreign Languages and Their Teaching*, (02), 10-16.

Li, J. (2007). On the Breakthrough of "Dumb English". *The Science Education Article Collects.* (07), 79,97.

Li, K. (2020). The Application of Task-based Teaching Method in College English Teaching. *English on Campus*, (29):1-3.

Li, S., & Fu, M. (2018). Strategic and Unpressured Within-task Planning and their Associations with Working Memory. *Language Teaching Research, 22*(2), 1-24.

Li, X. (2019). The Historical Process and Future Trend of College English Curriculum Construction in China. *Foreign Language Teaching and Research* (6), 814-818.

Littlewood, W. (1992). Cognitive Principles Underlying Task-centered Foreign

Language Learning. *Assignments*, 19.

Liu, C. (2020a). *A Psycholinguistic Study of Second Language Production.* Beijing: Science Press.

Liu, J. & Liu, L. (2017). *Fundamentals Application Course of SPSS Data Statistics and Analysis.* Beijing: Tsinghua University Press.

Liu, L. & Miu, H. (2018). Theory and Measurement of Language Complexity. *Foreign Languages Research,* (01),52-55.

Liu, Q. & Hu, Z. (2005). How to Use SPSS, SAS Statistical Software for Normality Test. *Journal of Xiangnan University*, 03, 60-62.

Liu, X. (2009). Research on the Current Situation and Countermeasures of Oral English Teaching for Science and Engineering. *China Electric Power Education* (11), 246-247.

Liu, X. (2010). Krashen's Input Hypothesis Theory and its Implications for College English Teaching. *Master*, (11), 137-138.

Liu, W. (2020b). Research on English Education in Modern China (1862-1937). Doctoral dissertation, Wuhan University.

Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. De Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspective*. Amsterdam: John Benjamins.

Long, M. (2015). *Second Language Acquisition and Task-Based Language Teaching*. Oxford: Wiley-Blackwell.

Long, M. & Robinson, P. (1998). Focus on Form: Theory, Research and Practice. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press.

Lu, Y. (2016). From Input to Output: An Analysis of Cognitive Factors Affecting L2 Oral Output. *Foreign Studies*, *004*(002), 37-42.

Luo, S. (2009). The Various Teaching Activities in the Pre-task Planning in Task-based Classroom Teaching. *Foreign Language Teaching in Schools (Middle Schools)* (01), 1-6.

Lynch, T. & Maclean, J. (2000). Exploring the Benefits of Task Repetition and Recycling for Classroom Language Learning. *Language Teaching Research*, *4*(3), 221-250.

McArthur, T. (2004). Is it World or International or Global English, and Does it Matter?. *English Today*, 20/ 3.

Mehnert, U. (1998). The Effects of Different Lengths of Time for Planning on Second Language Performance. *Studies in Second Language Acquisition*, *20*(01).

Ministry of Education (MOE). (2007). *Requirements for College English Teaching.* Beijing：Foreign Language Teaching and Research Press.

Ministry of Education (MOE). (2016). A Series of Higher Education Quality Reports Were Released for the First Time: Facts and Figures Show Self-confidence and Introspection of the Quality of Higher Education in China. Ministry of Education, April 7, 2016. Retrieved February 12th, 2023, from http://www.moe.gov.cn/jyb_xwfb/xw_fbh/moe_2069/xwfbh_2016n/xwfb_160407/160407_sfcl/201604/t20160406_236891.html

Ministry of Education (MOE). (2020). *General Guidelines of College English Teaching*. Beijing: Higher Education Press.

Ministry of Education (MOE). (2021). Grasping the Opportunity of "Double Reduction" to Improve English Classroom Quality. Ministry of Education, November 19, 2021. Retrieved February 12th, 2023, from http://www.moe.gov.cn/jyb_xwfb/moe_2082/2021/2021_zl53/zjwz/202111/t20211119_580961.html

Morgan, D. (2007). Paradigms Lost and Pragmatism Regained Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, 48-76.

Mu, H. (1988). A Study on the Standard Measurements of T-unit. *Foreign Language Teaching and Research.* (4), 52-58.

Nunan, D. (1989). *Designing Tasks for the Communicative Classroom.* Cambridge: Cambridge University Press.

Nunan, D. (2011). *Task-based Language Teaching*. Cambridge: Cambridge University Press.

Ortega, L. (1999). Planning and Focus on Form in L2 Oral Performance. *Studies in Second Language Acquisition, 21*(01), 109-148.

Ou, Y. & Huang, X. (2016). A Study on the Influence of Oral Task Types on English Majors' Cognitive Stance Expression. *Journal of Zhejiang International Studies University*, (001), 19-24.

Paris, D., & Reynolds, J. (1983). *The Logic of Policy Inquiry*. Longman.

Pattison, P. (1987). *Developing Communication Skills*. Cambridge: Cambridge University Press.

Patton, M. (1990). *Qualitative Evaluation and Research Methods (2nd Edition).* Newbury Park, CA: Sage.

Plowright, D. (2011). Ethical Issues in Participant-centred Research (Chp.12). *Using Mixed Methods: Frameworks for an Integrated Methodology*. London: Sage.

Polio, C. (2001). Research Methodology in Second Language Writing Research: The Case of Text Based Studies. In T. Silva & P. K. Matsuda (Eds.), *On Second Language Writing.* Lawrence Erlbaum.

Prabhu, N. (1983). Procedural Syllabuses. Paper presented at the RELC Seminar, Singapore.

Prabhu, N. (1987). *Second Language Pedagogy: A Perspective.* Oxford: Oxford University Press.

Qin, X. & Bi, J. (2015). *Quantitative Approaches and Quantitative Data Analysis in L2 Research.* Beijing: Foreign Language Teaching and Research Press.

Richards et al. (1985). *Longman Dictionary of Applied Linguistics*. London: Longman.

Richards, J., & Rodgers, T. (2001). *Approaches and Methods in Language Teaching*. Cambridge: Cambridge University Press.

Robinson, P. (1995). Attention, Memory, and the "Noticing" Hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (2002). *Individual Differences and Instructed Language Learning.*

Amsterdam: Benjamins.

Robinson, P. (2007). Aptitudes, Abilities, Contexts and Practice. In R. M. DeKeyser (Ed). *Practice in Second Language Learning: Perspectives from Applied Linguistics and Cognitive Psychology.* New York: Cambridge University Press.

Robson, C. (2011). *Real World Research: A Resource for Users of Social Research Methods in Applied Settings (3rd Edition).* Chichester: Wiley.

Samuda, V. & Bygate, M. (2010). *Tasks in Second Language Learning*. Beijing: Foreign Language Teaching and Research Press.

Samuda, V.& Madden, C. (1984). Task-based Test: Testing as a Reflection of Classroom Methodology. *Applied Linguistics*, (1):84-94.

Sawyer, M. & Ranta, L. (2001). Aptitude, Individual Differences, and Instructional Design. In P. Robinson (Ed.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Schmidt, R. (2012). Attention, Awareness, and Individual Differences in Language Learning. In W. M. Chan et al. (Eds.). *Perspectives on Individual Characteristics and Foreign Language Education*. Boston/Berlin: De Gruyter Mouton.

Shen, H. (2013). A Study on the Role Transition of College English Teachers from the Perspective of Students. *Education Forum, 000*(021), 34-35.

Sheng, D. (2010). The Application of Experiential Learning in College Oral English Teaching. *Education and Vocation, 000*(005), 142-143.

Sheng, H. (2015). On the New Direction of Hierarchical English Teaching Reform in Engineering Colleges and Universities: A Discussion from the “Guidelines to College English Teaching (Exposure Draft)”. *The Science Education Article Collects*, *000*(034), 164-165.

Skehan, P. (1989). *Individual Differences in Second Language Learning*. Edward Arnold.

Skehan, P. (1996). Second Language Acquisition Research and Task-based Instruction.

In J. Willis & D. Willis (Eds.) *Challenge and Change in Language Teaching*. Oxford: Heinemann.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skehan, P. (1999). *A Cognitive Approach to Language Learning.* Shanghai: Shanghai Foreign Language Education Press.

Skehan, P. (2003). Task-based Instruction. *Language Teaching,* (36)1,1-14.

Skehan, P. (2009). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied Linguistics, 30*(4), 510-532.

Skehan, P. (2011). *Researching Task: Performance, Assessment and Pedagogy*. Shanghai: Shanghai Foreign Language Education Press.

Skehan, P. (2014). *Processing Perspectives on Task Performance*. Amsterdam: John Benjamins Publishing Company.

Skehan, P., & Foster, P. (1997). Task Type and Task Processing Conditions as Influences on Foreign Language Performance. *Language Teaching Research, 1*(3), 185-211.

Skehan, P. & Foster, P. (1999). The Influence of Task Structure and Processing Conditions on Narrative Retellings. *Language Learning*, 49/1:93–120.

Skehan, P. & Foster, P. (2001). Cognition and Tasks. In P. Robinson (Ed). *Cognition and Second Language Instruction.* Cambridge: Cambridge University Press.

Skehan, P., & Shum, S. (2014). Structure and Processing Condition in Video-based Narrative Retelling. In P. Skehan. (Ed.). *Processing Perspectives on Task Performance*. Amsterdam: John Benjamins Publishing Company.

Stevick, E. (1989). *Success with Foreign Languages: Seven Who Achieved It and What Worked for Them.* London: Prentice Hall.

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques (2nd Edition).* Thousand Oaks, CA: Sage.

Su, R. (2012). A Study of Students' Expectation of Teachers' Role in Task-based Language Teaching. MA Dissertation, Central South University.

Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In S. Gass, & C. Madden (Eds.), *Input in Second Language Acquisition*. Rowley, MA: Newbury House.

Swain, M. (1995). Three Functions of Output in Second Language Learning. In G. Cook & B. Seidlhofer (Eds.), *Principles & Practice in Applied Linguistics*. Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1995). Problems in Output and the Cognitive Processes they Generate: A Step towards Second Language Learning. *Applied Linguistics, 16(3), 371-391.*

Tan, X. & Dong, R. (2007). The Influence of Task Type and Language Proficiency on Oral Accuracy and Complexity. *Journal of PLA University of Foreign Languages*, *030*(005), 54-58.

Tian, X. (2013a). The Implementation of Task-based Approach in College Oral English Teaching. *University Education* (09), 134-135.

Tian, Y. (2013b). The Extracurricular Support System of College Oral English from the Perspective of Input Hypothesis-Problems and Countermeasures: Taking College English Corner as an Example. *Journal of Language and Literature Studies*, *000*(007), 91-92.

Tong, S., Adamson, B., & Che, W. (2000). Tasks in English Language and Chinese Language. In B. Adamson, T. Kwan, and K. Chan (Eds.), *Changing the Curriculum: The Impact of Reform on Primary Schooling in Hong Kong.* Hong Kong: Hong Kong University Press.

Tyler, R. (2014). *The Basic Principles of Curriculum and Teaching*. Beijing: China Light Industry Press.

VanPatten, B. (1990). Attending to Form and Content in the Input: An Experiment in Consciousness. *Studies in Second Language Acquisition, 12*,287–301.

VanPatten, B. (2007). *From Input to Output: A Teacher's Guide to Second Language Acquisition*. Beijing: World Publishing Corporation.

Wang, G. (1999). *China's Population: Problems, Thoughts and Policies*. Aldershot:

Brookfield.

Wang, J. (2010). On the Process of Linguistic Capacity Constructing. *Foreign Language Research*, *000*(003), 88-91.

Wang, J. (2012). *SPSS Concise Operation Course-Case Analysis Oriented*. Hubei Science and Technology Press.

Wang, J., Gao, L. & Cheng, Y. (2013). A Study on the Effectiveness of Task-based Speech Training in College Oral English Teaching. *Journal of Hunan Institute of Posts and Telecommunications,12*(1), 102-105.

Wang, S. (2016). Interpretation of Key Points of College English Teaching Guide. *Foreign Language World,* (3), 2-10.

Wang, H. Z. & Song, F. Q. (2015). Effects of Strategic Planning Time on L2 Paired Oral Test Performance. *Chinese Journal of Applied Linguistics* (03),263-276,368.

Wang, S. & Wang, H. (2019). Keeping to the Right Direction and Promoting the Development of College English Teaching-A Summary and Reflection on the Work of College English Steering Committee in 2013-2017. *Foreign Language World, 191*(02), 9-15.

Wang, Y. & Wang, H. (2008). The Application of Functional Notional Approach in College Language Teaching. *Modern Education Management, 000*(001), 127-128.

Wang, Z. (2014). On-line Time Pressure Manipulations: L2 Speaking Performance under Five Types of Planning and Repetition Conditions. In P. Skehan. (Ed.). *Processing Perspectives on Task Performance*. Amsterdam: John Benjamins Publishing Company.

Wei, J. (2011). Task-driven Teaching Method: A Bridge from Literature Palace to Real Life. *Journal of Anhui Electronic Information Vocational and Technical College, 10*(6), 54-56.

Wen, Z. (2007). The Effect of Working Memory: A New Perspective of Foreign Language Learning Ability Research. *Modern Foreign Languages*, 87-95,110.

Willis, J. (1996). *A Framework for Task-based Learning*. Longman.

Wolfe-Quitero, K., Inagaki, K. & Kim, H. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu: University of Hawaii Press.

Wu, W. & Pan, K. M. (2012). Research on Task based Language Teaching in Asian Context. *Modern Foreign Languages* (03),312-318,330.

Wu, Y., Wu, G. & Ma, S, G. (2016). Review on the Origin, Genre, and Application of the Grounded Theory Method. *Journal of Distance Education*,*34*(3), 10.

Xia, J. (2001). The Benefits of Task-based Approach to College English Teaching. *China University Teaching*. (6), 32-34.

Xing, J. (2014). On the Three Dimensions of L2 Production: Complexity, Accuracy, and Fluency. *Journal of Beijing University of Chemical Technology (Social Science Edition)*, (4)

Xu, C., Zhang, J. & Zhan, J. (2017). Accuracy and Fluency in Chinese Learners' Writing in a Cross-Pacific Online Writing Course: A Dynamic Systems Perspective. *Foreign Languages in China*, *014*(006), 53-61.

Xu, H. (2015). The Influence of Task Structure and Preparation Conditions on Learners' Oral Output. *Foreign Languages and Their Teachings*, (1), 45-49.

Xu, J. & Chen, C. (2017). The Effects of Cognitive Requirements of Tasks on Learners' Language Output and Limited Attention Model and Multi-attention Resource Model. *Foreign Language Research*, (1), 91-96.

Xu, J. & Chen, C. (2018). The Effect of Cognitive Requirements on the Quality of Learners' Oral Output and the Distribution of Attention. *Foreign Languages and Their Teaching*, *000*(006), 42-52.

Yan, H, Guo, T., & Chen, W.Y. (2020). Analysis and Countermeasures of the "Three Teaching" Problems in Chinese Teaching in Senior High Schools in Thailand. *Educational Observation* (03), 98-100,130.

Yan, J. & Jia, H. (2006). English Speech and the Investigation of Students' Comprehensive Ability. *Journal of East China University of Science and Technology (Social Science Edition)* (01), 123-125.

Yang, G. & Xie, H. (2014). *From Novice to Expert: SPSS Data Statistics and Analysis.* Beijing: Tsinghua University Press.

Yang, W. (2014). Mapping the Relationships among the Cognitive Complexity of Independent Writing Tasks, L2 Writing Quality, and Complexity, Accuracy and Fluency of L2 Writing. *Dissertations & Theses-Gradworks*.

Yang, Y. (2015). *Psycholinguistics.* Beijing: Science Press.

Yao, X. (1993). Foreign Languages in China's Higher Education. *Language Learning Journal*, 7, 74-77.

Yi, B. (2014). Researches on Task Factors' Influence on Second Language Acquisition and Production. *Journal of Chengdu Normal University*,30(2), 6.

Young, R. (1995). Conversation Style in Language Proficiency Interview. *Language Learning*, 45/1, 3-45.

Yu. Q. M., Lu, X. H., & Sun, L. L. (2010). A Study of Register Features in Learner English Speeches. Journal of Changshu Institute of Technology(Philosophy & Social Sciences), (07),81-84.

Yu, T. & Lei, F. (2004). *Revolution in English Teaching: A New Book on Language Teaching Methods.* Lanzhou: Lanzhou University Press.

Yuan, F. & Ellis, R. (2003). The Effects of Pre-task Planning and On-line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. *Applied Linguistics, 24*(1), 1-27.

Zeng, W. (2005). Research on Task-based Teaching Method and its Application in China. *Journal of Basic English Education,7*(001), 34-39.

Zeng, W. (2007). A New Probe into the Core Issues of Task-based Language Teaching. *Teaching Administration,* (1), 3.

Zhang, H. (2014). Research and Practice of Oral English Teaching Strategies for Science and Engineering Students. *Teaching of Forestry Region,* (8), 66-67.

Zhang, P. (2011). Mixed Method Research: Paradigm, Design, and Quality Standard. *Journal of Tianjin University (Social Science Edition), 13*(005), 441-444.

Zhang, Q. (2019). *Language Production: Perspectives of Psycholinguistics.* Shanghai:

East China Normal University Press.

Zhang, Y. K. (2010). The Application of Task-based Teaching in College Oral English Teaching: A Case Analysis. *Journal of Qinghai University for Nationalities (Education Science Edition), 30*(001), 109-111.

Zhang, Y. P. (2006). The Theory and Practice of "Focus on Form" in Second Language Teaching. *Journal of Xi'an Foreign Studies University*, 14(4), 43-47.

Zheng, L. (2016). A Review of the Reform of College Oral English Teaching. *Journal of Beijing City University, No.132*(02), 43-50.

Zheng, M. (2010). The Role of College English Teachers from the Perspective of Students. *Journal of Chengdu Normal University*, 26(007), 7-9.

Zheng, Y. & Feng, Y. (2017). A Dynamic System Study on the Development of Learners' Syntax and Vocabulary Complexity. *Modern Foreign Languages*, (1), 57-68.

Zhong, L. L. (2021). The Influence of Task Complexity and Pre-task Planning on Oral Production of Senior High School Students. Master's Thesis, Jiangxi Normal University.

Zhou, L. & Li, Y. (2009). The Change of Pragmatism Philosophy-From Dewey to Hu Shi. *Vicissitudes*, (1), 164-165.

Zhou, M. (2019). The Development and Research of English Teaching in Modern China. *Shanxi Archives* (03),177-179

Zhou, X., Zhuang, J. & Shu, H. (2001) The Theoretical Framework of Speech Production Research. *Journal of Psychological Science* (03), 262-265.

Zhu, J. (2002). A War between Paradigms?-Debates about Quantitative and Qualitative Research: A Methodological Problem. *Studies in Dialectics of Nature, 018*(007), 22-24.

Zhu, S. (2007). The Application of Cooperative Principles in College Oral English Teaching. *Journal of Jiangxi Electric Power Vocational and Technical College*, (03), 31-32.

Zhu, W. (2022). The Efficient Application of Experiential Teaching Method in

College English Teaching. *Overseas English* (21), 138-140

Zhuo, Y. (2009). A Tentative Analysis on Independent Variable and Dependent Variable-The Development of Self-learning Ability and Creativity in College English Teaching. *Journal of Jilin Normal University (Humanities & Social Science Edition), 37*(004), 101-104.

Zou, J. (2014). Research on Embedding English Public Speaking into College English Class of Universities of Science and Technology. *Journal of Yunmeng*, (06), 139-142.

Zou, T. (1990). *John Dewey, A Pragmatist.* Jilin: Jilin Education Press.

Appendices

Appendix 1: Participant Consent Form

PARTICIPANT CONSENT FORM

Project title: Task and Language Production: A Study of Task Characteristics and Task Conditions on Spoken Language Production in English Speaking Course

Researcher's name: Ms. Xi Chen

Supervisor's name: Professor. Bob Adamson

- I have read the Participant Information Sheet and the nature and purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the research project at any stage and that this will not affect my status now or in the future.
- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that the interview/data collection [*omit as appropriate*].

will be recorded/filmed [*omit as appropriate*].

- I understand that data will be stored in accordance with data protection laws.

- I understand that I may contact the researcher or supervisor if I require more information about the research, and that I may contact the Research Ethics Sub-Committee of the University of Nottingham, Ningbo if I wish to make a complaint related to my involvement in the research.

Signed ……
(participant)

Print name …………

Date ……Oct 14th,2019……

Contact details

Researcher: Ms. Xi Chen and zx20360@nottingham.edu.cn

Supervisor: Professor Bob Adamson and bob.adamson@nottingham.edu.cn

UNNC Research Ethics Sub-Committee Coordinator:

Joanna.Huang@nottingham.edu.cn

Appendix 2: Sample Transcriptions of Students' Recordings

Phase 1 Students' Recordings

Label the students from A-F

Student A- SA

Student B -SB

Student C- SC

Student D- SD

Student E- SE

Student F- SF

Abbreviation

Fillers: ER; UH; URM

Pause: P

Long Pause: LP

Repetition: REPT

Reformation: REF

Replacement: REPL

P1 SA.mp3

With fluency features:

As a *[P]* as a *[REPT]* new transportation, *[ER]* shared bikes changed our lives greatly. *[ER]* But a new way for journey, it is convenient and cheap, and so on. *[ER]* But it *[REPL]* problems exist, too, [*REPT]*exist, but problems exist as either. The illegal parking *[ER]* The *[ER]* when that slims, *[ER]* the safet who [*REPT]* the safet who *[ER] [REPT]* who save the car? And the *[REPT]* the QR code were [*REPT]* the QR code were *[REPT]* were *[REPT] [UH]* were plagiitces *[ER]*, influence *[ER]*, change *[ER]*, our life. Yeah.

Without fluency features:

As a as a new transportation, shared bikes changed our lives greatly. But a new way for journey, it is convenient and cheap, and so on. But it problems exist, too, exist, but problems exist as either. The illegal parking The when that slims, the safet who the safet who who save the car? And the the QR code were the QR code were were were plagiitces, influence, change, our life. Yeah.

P1 SB.mp3

With fluency features:

Bike sharing is very convenient. *[P]* It is good for protecting the environment, and exercise. But sometimes, it's hard to find a bike, and *[UM]*, sometimes it is hard to find a place to park. Some bikes, *[ER] [REF]* some sharing, *[REF]* some bike is broken, so it is difficult to control and make all the sun *[REPL]* may cause danger. And, *[ER] [LP]* I think *[LP]* We should use it *[LP]* properly so that we will have a *[P]* better environment?

Without fluency features:

Bike sharing is very convenient. It is good for protecting the environment, and exercise. But sometimes, it's hard to find a bike, and sometimes it is hard to find a place to park. Some bikes, some sharing, some bike is broken, so it is difficult to control and make all the sun may cause danger. And, I think We should use it properly so that we will have a better environment?

P1 SC.mp3

With fluency features:

[ER] As for the influences of share bikes takes in our life, I think there is good, *[ER] [REF]* good ways *[REF]* good *[REF]* good influences or bad and bad influence. *[ER] [ER]* It is deliberately that the shared bike had changed our transportation and *[UH]* give us a *[ER]*another choice between *[ER]* so many *[UH]* choices. We can *[UH]* in *[UH]* whenever in whatever *[ER] [P]* whether long distance the journey or short distance journey. We all can choose a shared bike as our transportation. And it gives us many convenience.

Without fluency features:

As for the influences of share bikes takes in our life, I think there is good, good ways good good influences or bad and bad influence. It is deliberately that the shared bike had changed our transportation and give us a another choice between so many choices. We can in whenever in whatever whether long distance the journey or short distance journey. We all can choose a shared bike as our transportation. And it gives us many convenience.

P1 SD.mp3

With fluency features:

Recently, more and more shared bikes appeared in our lives. It *[ER][URM]* it has advantages and disadvantages. On the one hand, it is really makes our lives convenient. It is easily to get and you can see it everywhere. What you need is to own *[REF]* is to pay a little money and you can get it to everywhere you like. *[UH]* On the other hand, it also has some disadvantages. *[ER] [ER]* as if you need to scan the code, but you don't know if the bicycle is good or not. The bicycle you sign, *[REF]* you scan may be bad, so you need to scan another. So, we should use it. *[URM]*

Without fluency features:

Recently, more and more shared bikes appeared in our lives. It it has advantages and disadvantages. On the one hand, it is really makes our lives convenient. It is easily to get and you can see it everywhere. What you need is to own is to pay a little money and you can get it to everywhere you like. On the other hand, it also has some disadvantages. as if you need to scan the code, but you don't know if the bicycle is good or not. The bicycle you sign, you scan may be bad, so you need to scan another. So, we should use it.

P1 SE.mp3

With fluency features:

Shared bike huge *[REF]*, shared bike hugely changed our transportation, make our lifestyle more wonderful, bring convenience to our journey. Recent technology, developing *[P]* just brush the, QR cold, and you can enjoy it for a low place. Besides, GPS easily find it and lead you what to go. But, there are still some *[REF]* there are still some problems, which *[REPT]* which still to solve *[REPT]* solve, such as taking a real *[REF]* such as Peking areas is hard to find. Sometime, *[REF]* Sometimes I even *[P]* can not find it. And the bike damage is very important *[REF]* very important problems.

Without fluency features:

Shared bike huge, shared bike hugely changed our transportation, make our lifestyle more wonderful, bring convenience to our journey. Recent technology, developing just brush the, QR cold, and you can enjoy it for a low place. Besides, GPS easily find it and lead you what to go. But, there are still some there are still some problems, which which still to solve solve, such as taking a real such as Peking areas is hard to find. Sometime, Sometimes I even can not find it. And the bike damage is very important very important problems.

P1 SF.mp3

With fluency features:

Well, the transportation of shared bike is change our life, such as we can use shared bike to have a short *[REPT]* short *[REF]*short journey. *[ER]*When we need to use shared bike, we can take out our phone and then *[REF]* and then open the WeChat and scan the QR Code. And then the bill is paying. And the *[ER]* the shared bike is convenience and shaped to adult. We can use *[REF]* we can use WeChat to afford it. And, we can use shared bike everywhere in importing car porting laws. But there is a problem is that, we, it's hard to find shared bikes, because the people of *[REPT]* of user is too much. So, *[URM]* and Another problem is…

Without fluency features:

Well, the transportation of shared bike is change our life, such as we can use shared bike to have a short short short journey. When we need to use shared bike, we can take out our phone and then and then open the WeChat and scan the QR Code. And then the bill is paying. And the the shared bike is convenience and shaped to adult. We can use we can use WeChat to afford it. And, we can use shared bike everywhere in importing car porting laws. But there is a problem is that, we, it's hard to find shared bikes, because the people of of user is too much. So, and Another problem is…

Appendix 3: A Sample of Speech with Coding for CAF

P1 SA.mp3

With fluency features:

As a *[P]* as a *[REPT]* new transportation, *[ER]* shared bikes changed our lives greatly. *[ER]* But a new way for journey, it is convenient and cheap, and so on. *[ER]* But it *[REPL]* problems exist, too, [*REPT]*exist, but problems exist as either. The illegal parking *[ER]* The *[ER]* when that slims, *[ER]* the safet who [*REPT]* the safet who *[ER] [REPT]* who save the car? And the *[REPT]* the QR code were [*REPT]* the QR code were *[REPT]* were *[REPT] [UH]* were plagiitces *[ER]*, influence *[ER]*, change *[ER]*, our life. Yeah.

Without fluency features:

As a as a new transportation, shared bikes changed our lives greatly. But a new way for journey, it is convenient and cheap, and so on. But it problems exist, too, exist, but problems exist as either. The illegal parking The when that slims, the safet who the safet who who save the car? And the the QR code were the QR code were were were plagiitces, influence , change , our life. Yeah.

Phase 1	Concrete: a familiar topic	Unstructured: without the speech outline	Strategic planning	Students	SA
Complexity	**Lexical complexity:**				
	lexical sophistication: the proportion of low-frequency words in speaking;				5.56%
	lexical diversity: the ratio between parts of vocabulary (type) and token (TTR);				0.6
	lexical density: Content Word Ratio, CWR;				0.49
	Syntactic complexity:				
	the average length of AS-unit;				14.8
	clause per AS-unit.				2.6
Accuracy	the ratio of error-free clauses;				13.33
	the number of errors per hundred words: grammatical errors and pronounciation errors				0.00081
Fluency	**Temporal variables**:				
	speech rate: words per minute				74
	number of pauses and fillers: including pauses, long pauses and all the fillers				12
	Hesitation phenome repetitions				8
	reformations				0
	replacements				1

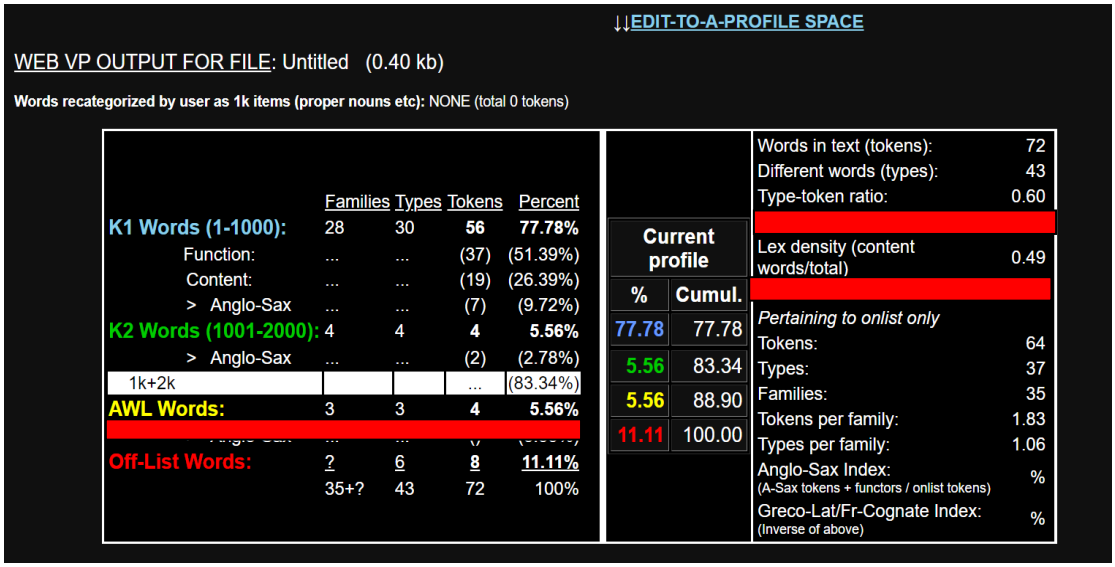
P1 SA

- **Lexical Complexity**

AWL Words: 5:56%

TTR: 0.6

Lexical density: 0.49



↓↓EDIT-TO-A-PROFILE SPACE

WEB VP OUTPUT FOR FILE: Untitled (0.40 kb)

Words recategorized by user as 1k items (proper nouns etc): NONE (total 0 tokens)

	Families	Types	Tokens	Percent
K1 Words (1-1000):	28	30	56	77.78%
Function:	...	...	(37)	(51.39%)
Content:	...	...	(19)	(26.39%)
> Anglo-Sax	...	...	(7)	(9.72%)
K2 Words (1001-2000):	4	4	4	5.56%
> Anglo-Sax	...	...	(2)	(2.78%)
1k+2k			...	(83.34%)
AWL Words:	3	3	4	5.56%
Off-List Words:	?	6	8	11.11%
	35+?	43	72	100%

Current profile	
%	Cumul.
77.78	77.78
5.56	83.34
5.56	88.90
11.11	100.00

Words in text (tokens):	72
Different words (types):	43
Type-token ratio:	0.60
Lex density (content words/total)	0.49
Pertaining to onlist only	
Tokens:	64
Types:	37
Families:	35
Tokens per family:	1.83
Types per family:	1.06
Anglo-Sax Index: (A-Sax tokens + functors / onlist tokens)	%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	%

- **Syntactic Complexity**

  The average length of AS unit: 74/5=14.8

  Clause per AS unit: 13/5=2.6

- **Accuracy:**

  The ratio of error-free clauses: 2/15*100=13.33

  The number of errors per hundred words: 6/74/100=0.00081

- **Fluency:**

  Speech rate: 74 words/minute

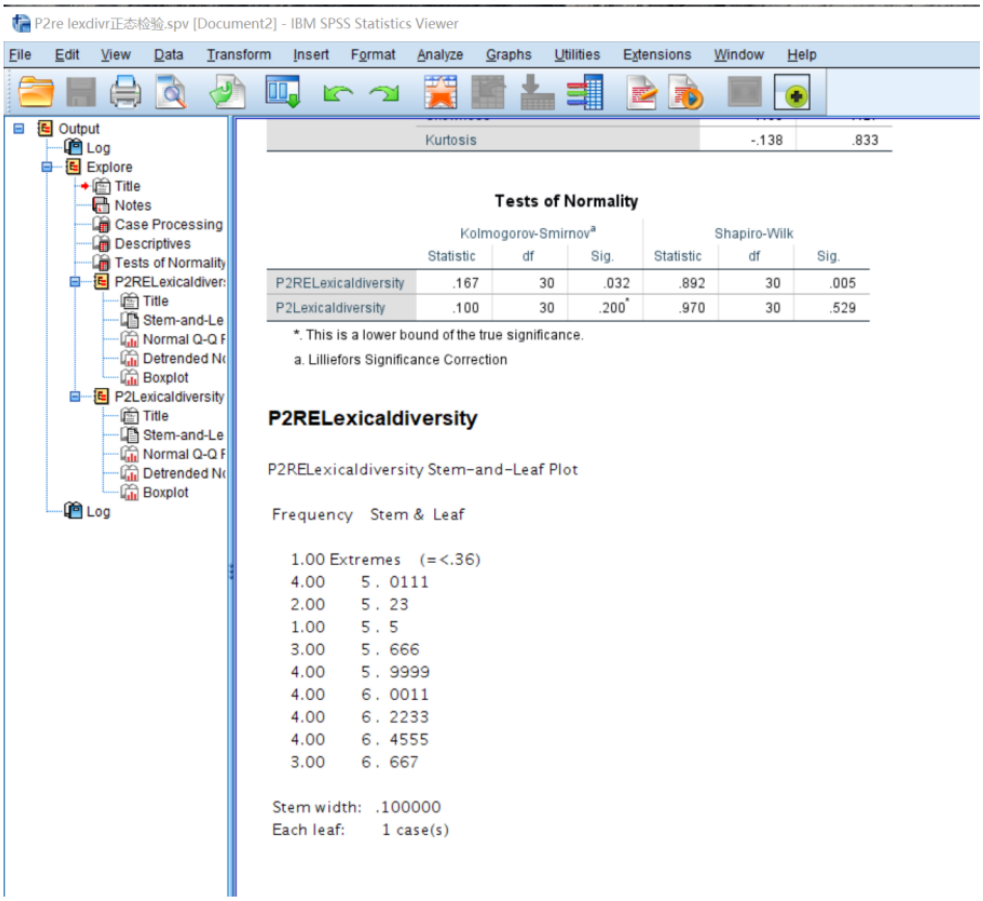
  Pauses and fillers: 12

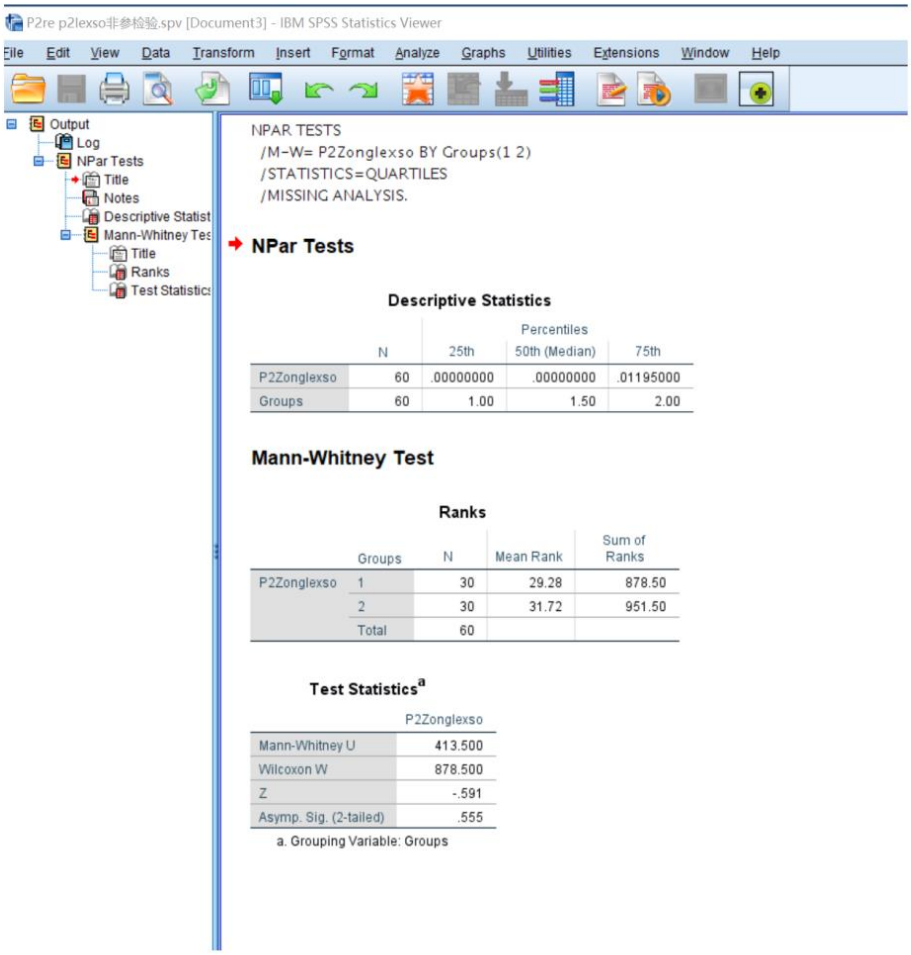
  Repetitions: 8

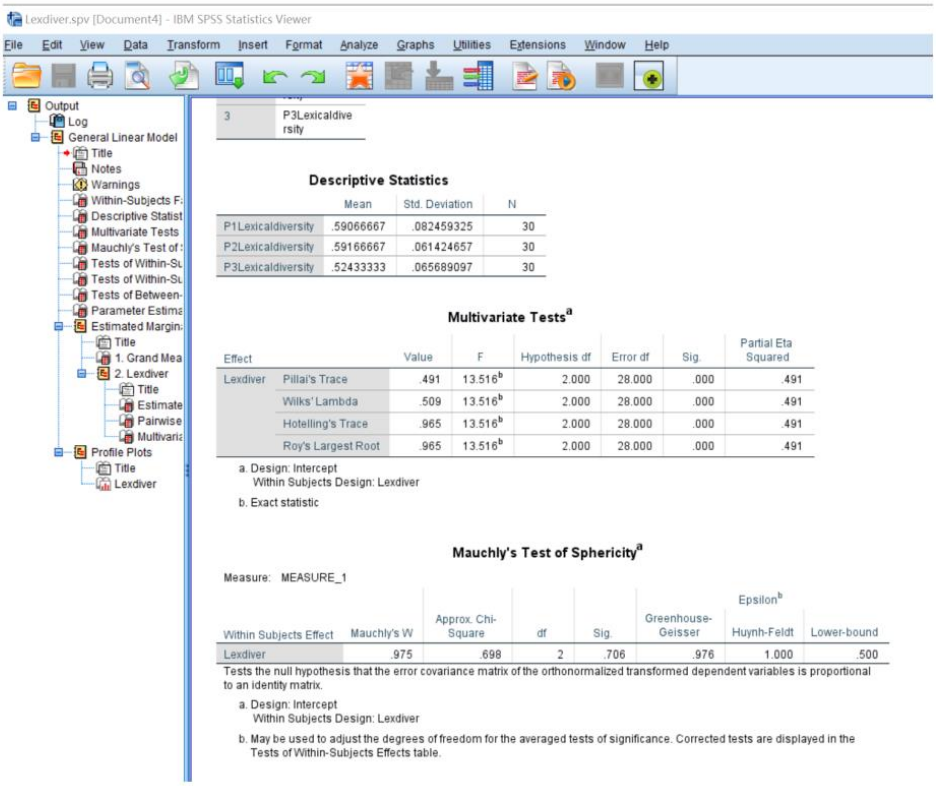
  Reformations: 0

  Replacements: 1

Appendix 4: Samples of SPSS Statistical Analysis




P2re p2lexso非參檢驗.spv [Document3] - IBM SPSS Statistics Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help
Output
Log
NPar Tests
Title
Notes
Descriptive Statist
Mann-Whitney Tes
Title
Ranks
Test Statistics
NPAR TESTS
/M-W= P2Zonglexso BY Groups(1 2)
/STATISTICS=QUARTILES
/MISSING ANALYSIS.
NPar Tests
Descriptive Statistics
N
Percentiles
25th
50th (Median)
75th
P2Zonglexso 60 .00000000 .00000000 .01195000
Groups 60 1.00 1.50 2.00
Mann-Whitney Test
Ranks
Groups N Mean Rank Sum of Ranks
P2Zonglexso 1 30 29.28 878.50
2 30 31.72 951.50
Total 60
Test Statistics[a]
P2Zonglexso
Mann-Whitney U 413.500
Wilcoxon W 878.500
Z -.591
Asymp. Sig. (2-tailed) .555
a. Grouping Variable: Groups



3	P3Lexicaldiversity

Descriptive Statistics

	Mean	Std. Deviation	N
P1Lexicaldiversity	.59066667	.082459325	30
P2Lexicaldiversity	.59166667	.061424657	30
P3Lexicaldiversity	.52433333	.065689097	30

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Lexdiver	Pillai's Trace	.491	13.516[b]	2.000	28.000	.000	.491
	Wilks' Lambda	.509	13.516[b]	2.000	28.000	.000	.491
	Hotelling's Trace	.965	13.516[b]	2.000	28.000	.000	.491
	Roy's Largest Root	.965	13.516[b]	2.000	28.000	.000	.491

a. Design: Intercept
Within Subjects Design: Lexdiver

b. Exact statistic

Mauchly's Test of Sphericity[a]

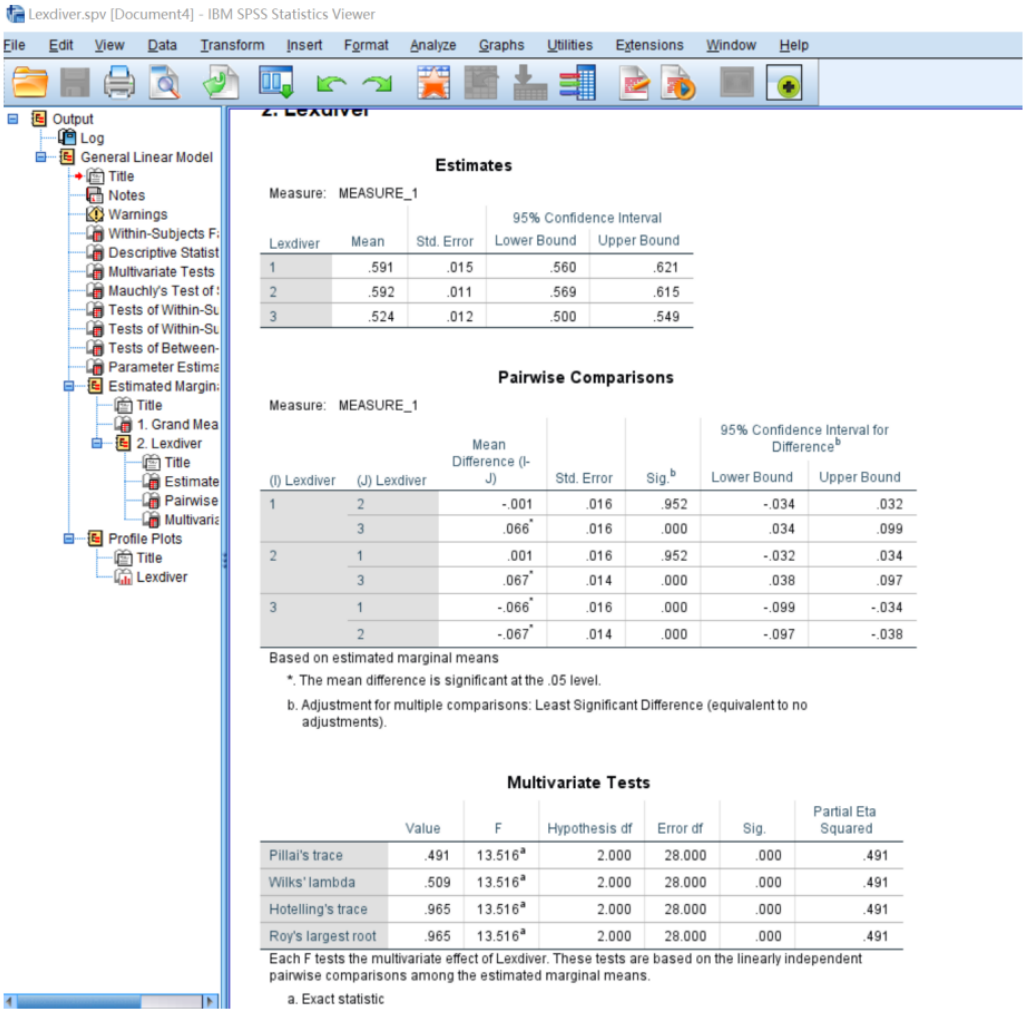
Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Lexdiver	.975	.698	2	.706	.976	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Lexdiver

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.



Estimates

Measure: MEASURE_1

Lexdiver	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	.591	.015	.560	.621
2	.592	.011	.569	.615
3	.524	.012	.500	.549

Pairwise Comparisons

Measure: MEASURE_1

(I) Lexdiver	(J) Lexdiver	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-.001	.016	.952	-.034	.032
	3	.066*	.016	.000	.034	.099
2	1	.001	.016	.952	-.032	.034
	3	.067*	.014	.000	.038	.097
3	1	-.066*	.016	.000	-.099	-.034
	2	-.067*	.014	.000	-.097	-.038

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.491	13.516[a]	2.000	28.000	.000	.491
Wilks' lambda	.509	13.516[a]	2.000	28.000	.000	.491
Hotelling's trace	.965	13.516[a]	2.000	28.000	.000	.491
Roy's largest root	.965	13.516[a]	2.000	28.000	.000	.491

Each F tests the multivariate effect of Lexdiver. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

Appendix 5: Sample Questionnaire Data

Task Design, Implementation and Spoken Language Production (II)

Dear students

Thank you so much for your time to participate in this questionnaire! The purpose of this questionnaire is to investigate the relationship between task design, implementation and spoken language production in English listening and speaking class, to analyze and explore the effective ways to improve the teaching oral output in this course.

Your authentic and serious answers play an important role in this study. Please complete the questionnaire according to your own specific experience in class. Your personal information and answers are anonymous and it is only for academic research purpose.

（A）Basic information

1. Gender: [Single choice]

Male	87	77.68%
Female	25	22.32%
Total	112	

2.Which range is your English score in the college entrance examination： [Single choice] *

Below 90	2	1.79%
90--110	0	0%
110--130	35	31.25%
130 above	75	66.96%
Total	112	

3. What is your major? [Single choice] *

Computer Science	37	33.04%
Architecture	11	9.82%
Automation	14	12.5%
Business Administration	8	7.14%
Energy and Power	3	2.68%
Civil Engineering	25	22.32%
Environmental Science and Engineering	14	12.5%
Total	112	

（B）Task design in oral English class

4. Which topic do you think is more familiar and easier for you to conduct oral speech in Week 2 and Week 3 of oral recording training? Please explain why. [Single choice] *

Describe a trip in Shenzhen	40	35.71%
Social networking pushes people closer or further	72	64.29%
Total	112	

5. Which topic do you think is more familiar and easier for you to conduct oral speech in Week 1 and Week 3 of oral recording training? Please explain why. [Single choice]

The influence of shared-bikes.	39	34.82%
Social networking pushes people closer or further	73	65.18%
Total	112	

6. Can more familiar topics help you produce more fluent, accurate, and complex language? Please explain why. [Single choice] *

Yes. It is helpful.	94	83.93%
Neutral.	17	15.18%
No, not helpful.	1	0.89%
Total	112	

7. In the first two weeks of the speech design, there is no explanation of the outline and structure of the speech. In the 3rd week, there are the explanation and preparation of the speech outline and structure.

Please compare the oral output of the first two weeks and the 3rd week. Will the 3rd week's presentation of outline and structure help you to produce more fluent, accurate, and complex language? Please explain why. [Single choice question] *

Yes. It is helpful.	88	78.57%
Neutral.	23	20.54%
No, not helpful.	1	0.89%
Total	112	

C）Task implementation in oral English class

8. For the task preparation before the speech, there are the different task preparation in the 2nd and 3rd week respectively. Which task preparation do you think is more helpful for your oral speech preparation to express more fluent, accurate, and complex language? Please explain why. [Single choice] *

Week 2：Rehearse your speech once	28	25%
Week 3：Peer discussion, brainstorming, model speeches, glossaries, rehearsal	84	75%
Total	112	

9. For the task preparation before the speech, there are the different task preparation in the second and third week respectively. Which task preparation do you think is more helpful for your oral speech preparation to express more fluent, accurate, and complex language? Please explain why. [Single choice] *

Week 1：Peer discussion, brainstorming, model	13	11.61%

speeches, glossaries		
Week 3：Peer discussion, brainstorming, model speeches, glossaries, rehearsal	99	88.39%
Total	112	

10. Based on the three weeks' recordings, which do you think is more fluent, accurate, and complex? Please explain why. [Single choice question] *

Week 1: The influence of shared-bikes on our life	19	16.96%
Week 2: Describe a travel experience in Shenzhen	28	25%
Week 3: Whether social networking pushes people closer or farther	65	58.04%
Total	112	

Thank you very much for your cooperation and support!

Appendix 6: A Sample Coding of the Open Question in the Questionnaire

9. For the task preparation before the speech, there are the different task preparation in the second and third week respectively. Which task preparation do you think is more helpful for your oral speech preparation to express more fluent, accurate, and complex language? Please explain why. [Single choice] *

Week 1：Peer discussion, brainstorming, model speeches, glossaries	13	11.61%
Week 3：Peer discussion, brainstorming, model speeches, glossaries, rehearsal	99	88.39%
Total	112	

Key words analysis:

Lexical frequency

Rehearsal 12

Better preparation 2

Fully prepared 2

Week 3 2

Speech 2

Topic 2

Helpful for correction 1

Reference after discussion 1

Outline 1

The more is better 1

Appendix 7: A Sample of Focus Group Interview Questions

Focus group interview 1: after phase 1 and 2

Theme 1: Task characteristics-familiarity of information; degree of structure

Familiarity of information

For topic 1 and topic 2 in the phase 1 and 2, which topic do you think you are more familiar with? Please explain why.

For the familiar topic, can it help to produce better language and why?

Theme 2-Task conditions-strategic planning and rehearsal

How do you think your speech produce when you are provided with strategic planning like: brainstorming for the speech contents; providing examples of model speeches and glossaries? Please explain why.

How do you think your speech produce when you can rehearse the whole speech? Please explain why.

Which helps you to produce language better? Strategic planning or rehearsal? Please explain why.

Appendix 8: A Sample of Focus Group Interview Transcript

Teacher: Last week, we have shared our views on the impact of shared-bicycles on our lives. Also, we recorded a speech on the travel experience in Shenzhen. I would like to ask you which topic you are more familiar with, and then could you explain the specific reasons for the two topics. Could each of us express our opinions in turn? So, last week, we had two topics, one was shared-bicycles, and one was Shenzhen travel experience. Which topic do you think will be more familiar?

Student 1: The shared-bicycles one.

Teacher: Why?

Student 1: Because I haven't been traveled in Shenzhen yet.

Teacher: Because everyone is busy studying, they haven't had time to go out, have they?

Student 1: It's mainly because I go home on National Day and don't bother to go out.

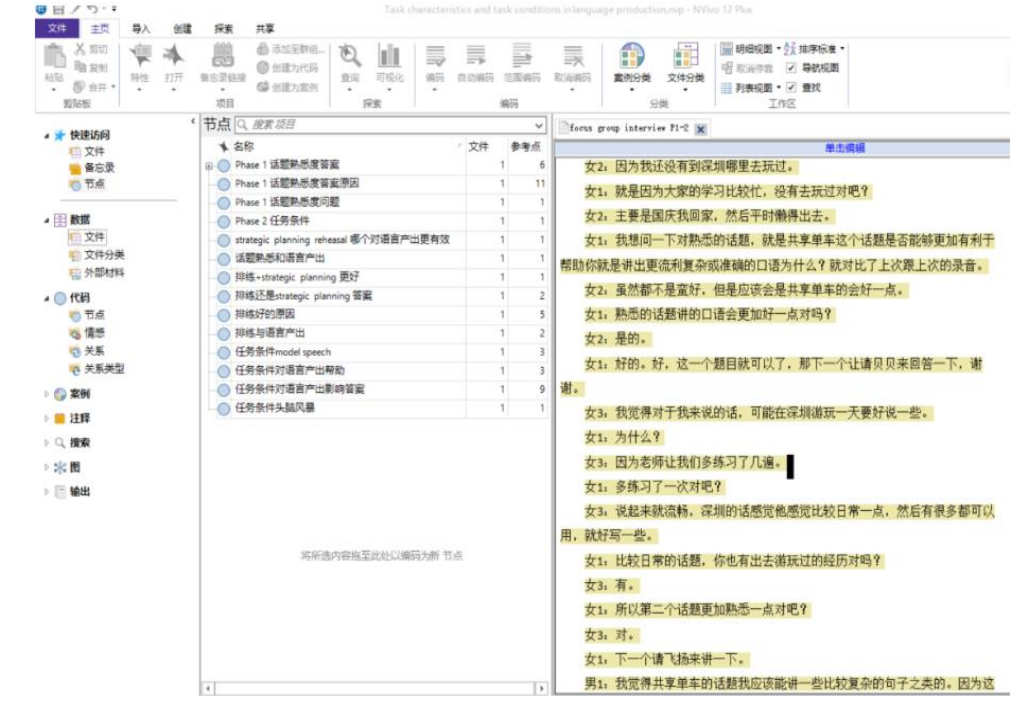
Teacher: I want to ask you about the familiar topic, that is, whether the “shared bicycles” topic can help you speak more fluently and accurately. Why?

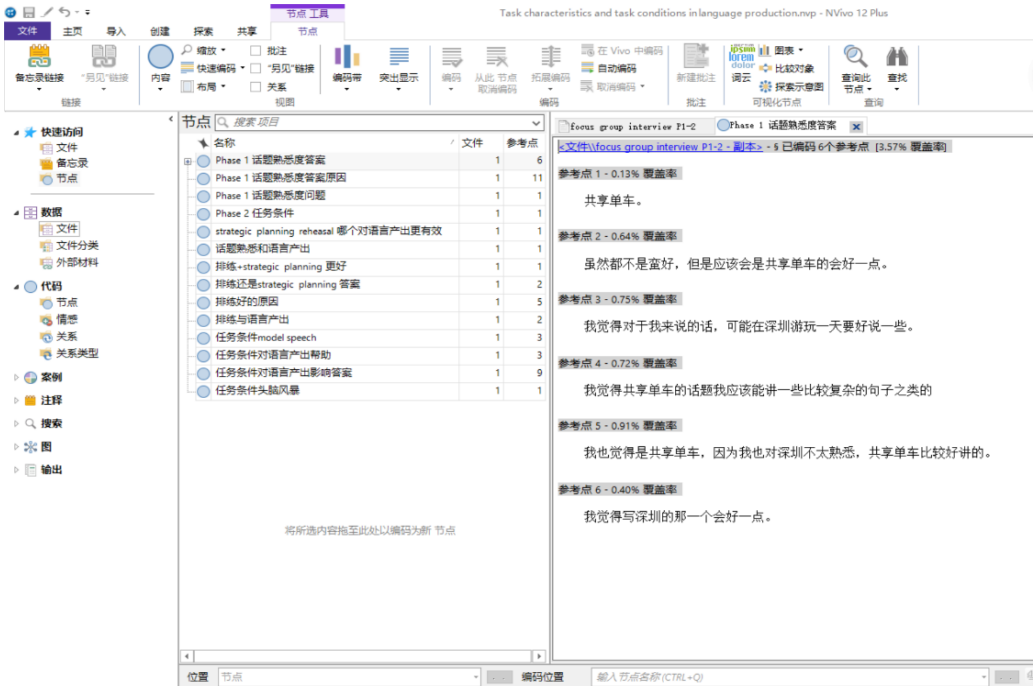
Student 1: Although they are not well spoken, it should be better for the “shared bicycles” topic.

Teacher: Speaking on familiar topics will be better, right?

Student 1: Yes.

Appendix 9: A Sample Coding of the Focus Group Data





Appendix 10: Detail Analysis of Themes and Nodes

Theme	File	Node
Action	0	0
Week 1.2.3 Topic Familiarity Comparison	1	2
Weeks 1 and 3 are familiar with the topic of shared-bikes and social networking	2	2
Topic familiarity	0	0
Topic familiarity is affected by the scope of the topic	2	2
Familiar topics are produced in spoken language	2	12
Week 1.2.3 Task Condition Comparison	1	3
Strategic planning and rehearsals	2	2
Compared Week 2 and Week 3 task conditions: Week 3 is better	2	10
Strategic planning and rehearsal in Week 3 were better.	2	2
The reason why the task conditions in Week 3 is better	2	10
Compare the Week 1 and Week 3 task conditions: Week 3 is better	2	2
The tasks in Week 1 and Week 3 are more well prepared	2	2
Rehearsal	0	0
The reasons for poor rehearsal	2	2
The reasons for good rehearsal	2	12
The rehearsal time is short and it is not easy to modify	2	2
Suggested rehearsal duration	2	2
Task implementation conditions help with speech-making	2	2

Outline of the speech	0	0
The speech outline doesn't help much	2	2
The speech outline needs to be well designed	2	2
The speech outline needs to be rehearsed once better	2	2
Reasons why the speech outline is useful	2	6
The speech outline is related to topic familiarity	2	6
Results of the spoken language production	0	0
Comparison of spoken language outputs in Week 1.2.3	0	0
The reasons why the oral output in Week 2 and 3 are not very good	2	2
The reasons why the Week 3 oral language output are good	2	12
The reason why the Week 1 speaking is good	2	4
Compare the spoken language output in Week 1.2.3	2	8
Compared to Week 1.2.3: Week 1 oral production is better	2	2
Conditions	0	0
Strategic planning	0	0
Glossary is good	2	14
Model speeches, examples and vocabulary are useful.	2	2
Model speeches are more useful than brainstorming	2	2
Model speeches are better	2	10
Model speeches limit ideas	2	2
Brainstorming and examples of shared-bikes are useful.	2	2
Brainstorming is not easy to operate.	2	6
It's better to rehearse once after brainstorming	2	2

Brainstorming is useful	2	12
Shared-bikes topic	0	0
The topic of shared-bikes is good.	2	22
Shared-bikes more fluent.	2	2
Shared-bikes understood better	2	6
Shared-bikes difficult to express	2	2
Wrote a shared-bikes essay	2	4
Rehearsal is good	2	12
Social networking topics	0	0
Students feel more of social networking than shared-bikes	2	4
The reasons why social networking is more familiar than shared-bikes	2	2
Students don’t know how to talk about social networking	2	2
People feel more deeply in social networking	2	2
The topic of social networking is more familiar	2	14
The reasons why social networking topic is more familiar	2	6
Shenzhen travel topic	0	0
Practice a few more times on the topic of travel in Shenzhen	2	2
The Shenzhen travel topic is more familiar	2	4
The topic of travel in Shenzhen is difficult to express	2	6
The topic of travel in Shenzhen is more daily for writing	2	2
The oral production of “travel in Shenzhen” is good	2	2
Students have few travel experiences in Shenzhen.	2	14
Students have travel experience in Shenzhen.	2	4
Students have made a speech on “travel in	2	4

Shenzhen”		
The speech outline is useful	2	8