การศึกษาผลกระทบของชนิดของกิจกรรมในแบบทดสอบการฟัง-การพูดด้วยคอมพิวเตอร์ที่มีต่อความสามารถทางการพูดของนักศึกษามหาวิทยาลัยราชภัฏบุรีรัมย์ชั้นปีที่ 1

นางสาวเสาวรจ เรืองไพศาล

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษเป็นภาษานานาชาติ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2552

THE EFFECTS OF COMPUTER-MEDIATED LISTENING-SPEAKING TEST TASKS ON THE ORAL ABILITIES OF FIRST-YEAR BURIRAM RAJABHAT UNIVERSITY STUDENTS

Ms. Saowarot Ruangpaisan

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in English as an International Language
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2009

Thesis Title	THE EFFECTS OF COMPUTER-MEDIATED LISTENING-SPEAKING TEST TASKS ON THE ORAL ABILITIES OF FIRST-YEAR BURIRAM RAJABHAT UNIVERSITY STUDENTS
By	Ms. Saowarot Ruangpaisan
Field of Study	English as an International Language
Thesis Advisor	Professor Kanchana Prapphal, Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

..... Dean of the Graduate School
(Associate Professor Pornpote Piumsomboon, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Assistant Professor Kulaporn Hiranburana, Ph.D.)

..... Thesis Advisor
(Professor Kanchana Prapphal, Ph.D.)

..... Examiner
(Professor Achara Wongsotorn, Ph.D.)

..... Examiner
(Assistant Professor Pavinee Thirakhupt, Ph.D.)

..... External Examiner
(Assistant Professor Wiwat Pantai, Ph.D.)

เสาวรจ เรืองไพศาล: การศึกษาผลกระทบของชนิดของกิจกรรมในแบบทดสอบการฟัง-การพูดด้วยคอมพิวเตอร์ ที่มีต่อความสามารถทางการพูดของนักศึกษามหาวิทยาลัยราชภัฏบุรีรัมย์ชั้นปีที่ 1. (THE EFECTS OF COMPUTER-MEDIATED LISTENING-SPEAKING TEST TASKS ON THE ORAL ABILITIES OF FIRST-YEAR BURIRAM RAJABHAT UNIVERSITY STUDENTS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ศ. ดร. กาญจนา ปราบพาล, 181 หน้า.

การวิจัยนี้มีจุดประสงค์เพื่อศึกษาผลของชนิดกิจกรรมในแบบทดสอบการฟัง – พูดด้วยคอมพิวเตอร์ 3 ประเภทคือ ตอบคำถาม (answering questions) บรรยายภาพ (describing pictures) และถ่ายโอนข้อมูล (transferring information) รวมทั้งระดับความสามารถทางภาษา 3 ระดับคือ ระดับความสามารถสูง (advanced) ปานกลาง (intermediate) และต่ำ(beginning) ที่มีต่อความสามารถทางการพูดของนักศึกษามหาวิทยาลัยราชภัฏบุรีรัมย์ชั้นปีที่ 1 กลุ่มตัวอย่างเป็นนักศึกษาชั้นปีที่ 1 วิชาเอกภาษาอังกฤษและภาษาอังกฤษธุรกิจจำนวน 84 คน ที่ได้รับการสุ่มแบบแบ่งชั้นตามคะแนนจากแบบทดสอบ The Cambridge Key English Test (KET) โดยแบ่งกลุ่มนักศึกษาตามความสามารถทางภาษา 3 ระดับ จากนั้นทดสอบด้วยแบบทดสอบการฟัง-การพูดด้วยคอมพิวเตอร์ซึ่งผู้วิจัยพัฒนาขึ้นเอง หลังจากนั้นคำตอบจะได้รับการประเมินโดยผู้เชี่ยวชาญ 2 ท่าน โดยประเมินความสามารถทางการพูดใน 4 ด้านได้แก่ ความถูกต้อง (accuracy) ความคล่องแคล่ว (fluency) ความซับซ้อน (complexity) และความเข้าใจ (comprehension) ข้อมูลที่ได้นำไปวิเคราะห์เชิงปริมาณโดยการวิเคราะห์ความแปรปรวนแบบสองทาง (two-way ANOVA) และการวิเคราะห์เนื้อหาเชิงคุณภาพ (Qualitative Analysis)

ผลการทดสอบพบว่าแบบทดสอบการฟัง - พูดทั้ง 3 ประเภทและระดับความสามารถทางภาษาของนักศึกษาทั้ง 3 ระดับ มีผลต่อความสามารถทางการพูดทั้ง 4 ด้านของนักศึกษาอย่างมีนัยสำคัญทางสถิติ อย่างไรก็ตามการวิจัยครั้งนี้ไม่พบปฏิสัมพันธ์ระหว่างระดับความสามารถทางภาษาและประเภทของแบบทดสอบการฟัง-พูดที่ส่งผลต่อความสามารถทางการพูดของนักศึกษา เมื่อพิจารณาความแตกต่างของความสามารถทางการพูดของนักศึกษาในแต่ละระดับพบว่า นักศึกษาที่มีระดับความสามารถทางภาษาในระดับสูงมีความสามารถทางการพูดทั้ง 4 ด้านสูงกว่านักศึกษาที่มีระดับความสามารถทางภาษาในระดับปานกลางและต่ำตามลำดับ และนักศึกษาที่มีระดับความสามารถทางภาษาในระดับปานกลางมีความ สามารถทางการพูดทั้ง 4 ด้านสูงกว่านักศึกษาที่มีระดับความสามารถทางภาษาในระดับต่ำอย่างมีนัยสำคัญทางสถิติ

เมื่อพิจารณาความแตกต่างของความสามารถทางการพูดทั้ง 4 ด้านของนักศึกษาจากการทำแบบทดสอบการฟัง-พูดทั้ง 3 ประเภทพบว่า คะแนนด้านความถูกต้อง ความซับซ้อนและความเข้าใจในภาษาจากการทำแบบทดสอบแบบตอบคำถามสูงกว่าแบบบรรยายภาพและถ่ายโอนข้อมูลอย่างมีนัยสำคัญทางสถิติ และคะแนนด้านความเข้าใจจากการทำแบบทดสอบแบบถ่ายโอนข้อมูล สูงกว่าคะแนนจากแบบบรรยายภาพอย่างมีนัยสำคัญทางสถิติ นอกจากนั้นคะแนนด้านความคล่องแคล่วในการใช้ภาษาจากการทำแบบทดสอบแบบตอบคำถาม สูงกว่าคะแนนจากแบบบรรยายภาพอย่างมีนัยสำคัญทางสถิติ ผลจากการวิเคราะห์ข้อมูลเชิงปริมาณสามารถนำไปใช้ในการออกแบบสื่อการสอน ตลอดจนการวัดผลการเรียนในวิชาการฟังและการพูด เพื่อพัฒนาความสามารถทางการพูดทั้ง 4 ด้านอย่างสมดุลย์ ผลจากการวิเคราะห์เนื้อหาเชิงคุณภาพให้ข้อมูลที่เป็นประโยชน์ต่อหลักสูตรการเตรียมความพร้อม และการจัดการสอนเสริมทางด้านการพูดสำหรับผู้เรียนที่มีความสามารถแตกต่างกัน

สาขาวิชา ภาษาอังกฤษเป็นภาษานานาชาติ ลายมือชื่อนิสิต.....

ปีการศึกษา 2552 ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

4889684920: MAJOR: ENGLISH AS AN INTERNATIONAL LANGUAGE
KEYWORDS: COMPUTER-MEDIATED LISTENING-SPEAKING TEST TASKS / ORAL ABILITIES / FIRST-YEAR BURIRAM RAJABHAT UNIVERSITY STUDENTS

SAOWAROT RUANGPAISAN : THE EFFECTS OF COMPUTER-MEDIATED LISTENING-SPEAKING TEST TASKS ON THE ORAL ABILITIES OF FIRST YEAR BURIRAM RAJABHAT UNIVERSITY STUDENTS. THESIS ADVISOR: PROF.KANCHANA PRAPPHAL, PH.D., 181 pp.

This study investigated the effects of the computer-mediated listening-speaking test consisting of 3 tasks: answering questions, describing pictures, and transferring information on the oral abilities of first-year Buriram Rajabhat University students. The study was conducted on 84 English and Business English major students. The Cambridge Key English Test (KET) was used to classify the students into three levels (advanced, intermediate and beginning) according to their levels of language abilities. The subjects were selected by the stratified random sampling method. Then, they took the computer-mediated listening-speaking test tasks developed by the researcher. Their oral responses were scored by two raters in terms of accuracy, fluency, complexity, and comprehension. Two-way Analysis of Variance (ANOVA) and descriptive statistics were used to analyze and interpret the data. In addition, the data were investigated qualitatively by using content analysis.

The results from the test tasks reveal that the three levels of the language abilities (advanced, intermediate and beginning) and the three test task types, namely, the answering questions, describing pictures and transferring information tasks significantly affected the oral abilities in the four language components. However, the interaction effect between the levels of the language abilities and task types was not found.

With regard to the four language components of the oral abilities of the three levels of the students, it was found that the advanced students gained significantly higher scores than the other two groups. Moreover, the intermediate group got higher scores than the beginning one.

Regarding the four language components of the oral abilities of the students performing the three task types, the accuracy, complexity and comprehension scores of the answering questions task were significantly higher than those from the describing pictures and transferring information tasks. In addition, the comprehension scores of the transferring information task were significantly higher than those from the describing pictures task. As for the fluency scores, they were higher in the answering questions task than in the describing pictures task. The quantitative data can be used to design instructional materials, and the assessment of listening and speaking courses with the balanced development of all four language components. The results from content analysis provide useful information to the preparation or remedial oral courses for diverse students.

Field of Study : English as an International Language Student's Signature Saowarot Ruangpaisa

Academic Year : 2009 Advisor's Signature Kanchana Prapphal

ACKNOWLEDGEMENTS

This study has been made possible through the invaluable suggestions and support given to me by my advisor, Professor Dr. Kanchana Phapphal, who was available whenever I needed her advice and encouragement. My great appreciation also goes to other members of my dissertation committee: Asst. Prof. Dr. Kulaporn Hiranburana, Prof. Dr. Achara Wongsotorn, Asst. Prof. Dr. Pavinee Thirakhupt, and Asst. Prof. Dr. Wiwat Puntai who served as the external examiner.

I would like to express my gratitude to Assoc. Prof. Dr. Suphat Sukamolson, Asst. Prof. Dr. Kulaya Benjakarn, Asst. Prof. Anu Yodprommin, Ajarn Dr. Sutthirak Sapsirin and Mr. Ivan Benedict New for their timely, useful and insightful comments for validating my instruments. I also acknowledge Ajarn Onuma Lakarnchua's reading of an early draft and the effort to help me correct some of the errors. Thanks also to Khun Nattawut Tartong for writing the computer software program for my computer-mediated test .

I want to express my deepest thanks to my departmental colleagues who substituted for my teaching loads and to Ms. Lilibeth Lago, Ajarn Chutima Bunrang, and Ajarn Atikom Kongdee who helped me in the instruments development and the data collection. My thanks also go to all students at Buriram Rajabhat University who willingly cooperated in this study. I would like to thank Assoc. Prof. Dr. Wilailak Kirin, Dr. Chatpramon Putichan, Ajarn Wongjan Poolpoem, and EIL colleagues who listened to me and encouraged me to overcome the obstacles during my study.

I am deeply grateful to Buriram Rajabhat University for the financial support for my tuition fee and the Graduate School, Chulalongkorn University for the scholarship for my research.

Last but not least, I am indebted to my family, especially my parents, Sanong and Sodsuk, who have been my constant thought of inspiration for my doctoral study and my sources of support, financially and emotionally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I

INTRODUCTION

1.1 Background of the Study

Though modern language test development has been of interest in 1920s (Spolsky, 1995), little attention has been paid to second language speaking tests due to the problems of reliability in scoring and practicality in test administration. However, due to US political and military needs, the first speaking test was developed in 1930 to screen overseas students applying to US colleges and universities (Fulcher, 2003). Until the present time, the test design has influenced “US immigration policy” and the political impact has also affected test development to consider immigration regulation and asylum seekers (Shohamy, 2001 cited in Fulcher, 2003). Indirect methods of testing speaking - that is, via other language skills or testing only sub-language components such as pronunciation and vocabulary, were used for centuries. There were attempts to have individual speaking tests but it could be only in the form of paper and pencil tests due to the subjectivity in scoring and logical administration.

It was not until the Second World War when the focus of language testing moved to speaking and communicative testing. As it was found that American diplomats and service personnel lacked language skills to perform their duties effectively, the Army Specialized Training Program (ASTP) was set to be the first language instruction program focusing on speaking second languages in specific areas (Fulcher, 1998). The test responding to this program was produced at Queen’s College, New York and it was later developed to be the first published test of speaking, the Foreign Service Institute (FSI) Oral Proficiency Interview (OPI) (Fulcher, 1998; 2000; 2003). Early speaking tests were influenced by the politics and military with the purpose of screening the immigrants to the US and to assure the effective duties of American service personnel. It can be said that the Second World War was the most crucial period for the beginning of modern speaking tests.

In 1956 the FSI speaking test first included interview techniques that allowed a test taker to converse with trained raters. The Educational Testing Service (ETS) also developed analytical rating scales to identify the students' speaking abilities more explicitly and fairly than the holistic criteria used before. However, later in the 1960s, due to the expanding US educational system, a large number of multiple choices item tests together with automatic marking machines were produced to serve the increasing needs of tests on industrial scale testing. But the scores of multiple choice test items could not accurately reflect students' ability of language use as these types of tests can measure only passive knowledge not the dynamic abilities (Fulcher, 1998). Due to the accurate measurement, the FSI's new testing procedures gained greater popularity and were implemented by numerous agencies in the United States, namely, CIA, the Peace Corps and the Defense Language Institute. In 1968 these US testing agencies developed a standardized testing system called the Interagency Language Roundtable (ILR). Test development process in 1950s to 1960s shed light on the more appropriate ways to rate speaking ability in order to avoid test bias and increase test reliability. These developments became the models for later researches (Fulcher, 1998; 2000; 2003).

In the 1970s the FSI system was introduced in non-government settings. Its use spread to many universities and states for certifying bilingual teachers. The OPI was well accepted because of not only its high face validity as a direct test of speaking ability and constant inter-rater reliability but also its support for the new notional-functional language teaching approaches which were very popular at that time. In 1979, the finding of Carrol's study (1967 cited in Fulcher, 2003) was of interest in that the scores on the FSI rating scales (going from 1 to 6) of college students majoring in foreign languages were only at level 2 or level 2 with plus point ($2/2^{+}$) which were under the minimum for professional working proficiency (FSI level 3). Following the study by Carrol, the Educational Testing Service (ETS) suggested that it does not mean that the students at this level ($2/2^{+}$) did not perform any progress after studying second languages. Thus, the ILR rating scales were developed to explain more about the students who were below this level. In 1979, a National Criteria and Assessment Program was set up and done by the American Council on the Teaching of Foreign Languages (ACTFL) to produce language tests and assess language learning in the US. ACTFL also increased score discrimination by adding a

number of bands (levels) at the lower end of the scale. ETS and ACTFL played a part in developing the ILR rating scales to respond to the purpose of speaking test in non-government setting. The first “ACTFL Provisional Proficiency Guidelines”, the final and the revision ones were published in 1982, 1986 and 1999 respectively and these have been the guidelines for the proficiency movement up until today (Fulcher, 1998; 2000).

Apart from the development of testing speaking administered by the formal organizations in America, educators have diverse perspectives regarding how to assess speaking abilities. McNamara (2000), for example, argues that instead of traditional discrete-point tests, more attention should be paid to a language test which involves integrated language performance. With the purpose of selecting foreign students for universities in the USA and UK in the 1990s, this kind of test could assess productive language abilities. Testing speaking should utilize more productive test types such as oral interviews and presentations with more emphasis on the linguistic knowledge over language use. Following Underhill (2003: 3), assessing speaking requires “ a genuine oral test that real people meet face to face and talk to each other ”. Hughes (2002) supports the idea that rather than assessing the mechanics of the interaction, the success in communication should rely on the ideas, emotion and information the interlocutors intend to convey. Similarly, Weir (2005) proposes the idea that in assessing oral ability, the learners should participate in “direct spoken language activities”; that is, they need to communicate in language tasks and adjust their speech to suit the unexpected situations, respond with fluency, and make good decisions under time constraints. Therefore, good speaking tests should allow test takers to communicate directly to the interlocutors and their achievements are not decided from fixed answers, but their language abilities in expressing themselves appropriately in a conversation.

According to Luoma (2004: 96), in designing a speaking test, test developers need to choose theoretical models that are related to the language abilities of test takers they would like to explain. Apart from models of language learning, the test construct can follow learning and teaching objectives. It is slightly more difficult to design the construct of the test based on teaching curricula since they are not presented explicitly in textbooks or any other “curriculum documents”. With the

"explicit concise format", models of language learning include the constructs which are easy to follow.

Among numerous models of language learning, communicative competence has been popular in language learning and assessment. Luoma (2004) states that communicative competence emphasizes not only the users but their language uses as well. The communicative competence models that have been used until now are based on the theory of language use in social life proposed by Hymes (1971, 1972). His theory includes understanding of language and how to use language in communicative context. Though his theory was discovered in the early of 1970s, it took almost a decade to incorporate Hymes's theory into the development of communicative language testing. McNamara (2000), for instance, proposed two features of communicative language tests based on Hymes' theory as follows:

1. the performance test that assesses learners' receptive and/or productive skills in communication, and

2. the test that is designed on the participants' social roles

According to Morrow (1979), the following three characteristics need to be considered whether the tests are communicative tests :

1. Communicative tests involve performance.

Test takers have to participate in real oral interactions by expressing the idea and using integrated skills. Fulcher (2000) suggests that the method of assessment should assure that the test performance is similar to the criterion performance. Thus, the test tasks and target language use tasks should be closely related. Most performance-based tests involve some situations such as role plays or simulations in which the learners have to assume roles in given situations; for example, taking the patients' information, giving lectures or engaging in business transactions. Speaking performance will be assessed "subjectively, qualitatively and impressionistically" and the scores on the performance test will show the efficiency of learners' performances.

2. Communicative tests are authentic.

Communicative tests are authentic in terms of their purposes, inputs and prompts as well as the contexts. The examiners have to recognize the communicative purposes of the tasks in order to respond to the tests appropriately. The input in the test should not be simplified. The test takers should be tested in the situations in which they have to select appropriate language for real contexts.

3. Communicative tests are scored on real-life or behavioral outcomes.

Success in a communicative test is determined on real-life or behavioral outcomes decided from whether the test takers perform the tasks by achieving their communicative goals. In the performance tests, test taker's performance is assessed on rating scales that show whether they successfully achieved a satisfactory outcome.

With the widespread implementation of Communicative Language Teaching (CLT) in the educational policies in East Asia for more than two decades (Littlewood, 2007), CLT has extended to Task-based Language Teaching (TBLT). Both CLT and TBLT share the same purpose of developing students' communicative competence (Richards, 2005). TBLT can be characterized as the latest version of CLT and a policy in the educational systems that the governments in the Asian Pacific region emphasize on. If CLT is the educational philosophy on which the language curricula are based, TBLT will be a sub-component in that system, namely, a teaching syllabus or methodology (Nunan, 2003; 2004). Tasks have been defined as the real-world activities learners try to accomplish by using the target language and focusing on meaning (Bachman and Palmer, 1996; Willis, 1996; Williams and Burden, 1997; Bygate Skehan, and Swain, 2001). The concept of tasks has been accepted among educators as a crucial element in language pedagogy, learning materials and syllabus design. Implementing Task-based Language Teaching (TBLT) helps improve learners' awareness of the target language and fluency in language use. Numerous educators view tasks as tools to promote second language acquisition. From a study by Morris (1994 cited in Meechai Iemjinda, 2004), after the TBLT was implemented with primary schools students in Hong Kong, they expressed their enthusiasm and willingness in learning English by using TBLT. For instance, English as a Foreign Language (EFL) countries, such as Hong Kong, China and Korea have integrated tasks into their educational policies (Prabhu, 1987; Long and Crookes, 1993; Nunan, 2004). Even in some countries like the Philippines, Indonesia, Japan and Thailand, though TBLT is not officially included in the syllabus, the concept is integrated into language activities (Vilches, 2003; Mukminatien, 2004; Viernan, 2005; Todd, 2006 cited in Littlewood, 2007).

Apart from being used as tools to aid learners' language learning, tasks are also used in assessing students' L2 communication abilities. In task-based assessment, tasks are used as tools for eliciting and evaluating learners' communicative performance or as authentic samples of target language requiring learners to perform real-life tasks which are meaning-focused and directed toward some specific roles

(Long and Norris, 2000; Elder, Iwashita and McNamara, 2002; Ellis, 2003). In addition, task-based assessment closely matches the test performance with the criterion performance (what the test takers have to do in the real world). Furthermore, using tasks as tools provides numerous benefits: it gives a washback effect; it helps integrate assessment into the learning process easily; it proposes the useful feedback about learners' strengths and weaknesses; and it provides test results which are easy to understand (Ellis, 2003).

When using tasks in eliciting language abilities, some researchers pay more attention to the following two factors: the features of the tasks or "task characteristics" and the conditions in which the tasks are administered or task conditions in way that they affect the tasks and the test takers' oral performance. Skehan (2001), for example, studied the effect of five task characteristics (familiarity of information, dialogic and monologic tasks, degree of structure, complexity of outcome and transformations) on examinees' language performance. Though he found no significant effects, his research results shed light on variables that other researchers should take into consideration. Following Skehan's idea, Wiggleworth (2001) conducted a study of the effects of two test characteristics (structure / non-structure and task familiarity) as well as two task conditions (native and non-native speakers (NS / NNS) and planning time) on learners' language performances. Similar to the result from Skehan (2001), there were no significant differences in learners' performances. Elder et al. (2001) also studied the impact of the manipulation of task instructions, materials and performance conditions on the fluency, accuracy and complexity of test takers' speech. They investigated three task design variables in picture narration tasks in semi-direct speaking tasks, i.e. immediacy (telling a story with or without picture clues); adequacy (telling a story from a set of 6 pictures or a set of 5 pictures with 1 picture missing); perspective (telling a story as it happened to you (first person) or from someone's point of view (third person)). However, no significant effect was found for any of the measures except "telling a story without picture clues" which produced more accurate language use.

In addition to the influence of task conditions and task characteristics on the oral abilities, other researchers especially from the EFL countries put the emphasis on the effect of task types. Teng (2007) investigated the effect of three different task types: answering questions, picture description and presentation on Taiwanese college students' speaking performance. Though the students' scores were not significantly

different on holistic measures, the analytical differences in terms of three language components (accuracy, complexity and fluency) were found. The complexity scores of the answering questions task is higher than those of picture description. Moreover, the students spoke more fluently in answering the questions task than in the other two tasks. Weir and Wu (2006) investigated the task comparability of three forms of the General English Proficiency Test Intermediate Speaking Test (GEPTS-I) used with intermediate EFL Taiwanese students. Each form contains three different task types: read-aloud, answering questions and picture description. The three tasks are statistically parallel and Test Form Two and Form Three are parallel at the individual task level.

As mentioned earlier, factors concerning tasks, namely task characteristics, task conditions and task types have been studied in order to design the appropriate formats of language tests. Furthermore, the ways to propose the tests interest many researchers. In order to catch up with the advances in information technology, the use of IT for developing computer-assisted language tests has existed and continued to increase (Brown, 1997; Alderson, 2000). Therefore, speaking should not be limited to communication between interlocutors. Interacting with computers or with other people via computers has become common for those concerned with language learning and testing. Assessing speaking through computers, then, is presented as an alternative of the typical speaking test types like paper-based tests, interviews or oral presentations and an alternative to language learners who would like to try simulations to prepare themselves before facing a real situation (Chapelle 2001). Further, the test takers can be presented with the multimedia in a more realistic way than paper-based tests. The advantage lies in the fact that this alternative lends itself to simulate reality in a language testing context and it shows relevant aspects of the target language use domain more accurately. Thus, the computer-based test helps increase authenticity and validity of the test (Choi, Kim, and Boo, 2003; Douglas and Hegelheimer, 2007). It is also more convenient to deliver the test via computers. For test administration, students can work individually anywhere and anytime without human proctors. In terms of scoring process, computers can report the scores immediately, more accurately and reliably and give instant diagnostic feedback to the test takers (Brown, 1997; Davidson, 2004). Delivering tests via computers, according to Alderson (2000) is like linking testing and assessment to teaching and learning in that the learners can check their progress after learning via the computer-assisted test

which is relied on the teaching and learning objectives. Computers will be more useful in all stages of test administration, e.g. test construction, test delivery, response analysis and a score reporting (Brown, 1997), especially for low-stakes testing. Moreover, the test takers have positive attitudes toward computer-based tests (Brown, 1997; Sticker, Wilder, and Rock, 2004; Teng, 2007).

Buriram Rajabhat University (BRU) has provided "Communicative English I and II courses" for English-major students but it has been found that they cannot communicate well in their ensueing careers. It may be because ways of teaching and assessment do not serve the demand for English uses. Assessing speaking via the traditional test methods such as multiple choices or short answers does not indicate their actual oral abilities. In addition, they lack opportunities to practice English in their daily lives. As a matter of fact, the students really need new methods to directly assess their speaking performances by providing the situation that is close to the real situation to compensate for the deficiency of enthusiasm, confidence in using languages and opportunities to be exposed to the target language. Despite its effectiveness, a direct speaking test is impractical because of the time constraint, a limited budget and lack of administrators. A semi-direct speaking test, however, can better address these limitations. Since the direct way to assess oral abilities like the face-to-face interview is too costly and impractical for the number of first-year English major students, a semi-direct speaking test task will be more appropriate for this situation. This study; therefore, aims to investigate the effects of task types on students' language performance in order to elucidate guidelines for designing a test which is suitable for the learners' language proficiency and is able to assess their aspects of language performance.

In addition, as very little research has been conducted on computer - mediated language tests of oral ability, there is little information about them. Computer-mediated listening-speaking test tasks are needed to provide beneficial language functions for students to enhance their oral abilities and encourage them to interact with more confidence. It can also help teachers to formulate teaching objectives, use teaching methods and design materials to facilitate learners' speaking skill. Besides, the Computer-mediated Listening-Speaking Test Task (CMLSTT) simulates real-life settings in language testing context before the students face the real situation. Furthermore, the English Program of Buriram Rajabhat University will have a listening-speaking test that can be administered to a large number of students. With

this newly-developed listening-speaking test, their oral abilities will be appropriately evaluated and a suitable syllabus can be developed based on the abilities of students at different levels.

1.2 Objectives of the Study

The objectives of the study are to

1. compare the oral abilities of the three groups of students in terms of different task types,
2. investigate the effects of 3 test task types of CMLST and three groups of students on the oral abilities,
3. study the differences in the accuracy, fluency, complexity, and comprehension of the test-takers' oral performances in each task type.

1.3 Research Questions

1. Do the advanced, intermediate and beginning students perform differently on three types of listening-speaking test tasks?
2. Do the different test task types and groups of students affect their oral abilities?
3. How do each test-takers' oral ability differ in their accuracy, fluency, complexity, and comprehension as measured by each test task type?

1.4 Statement of Hypotheses

Hypothesis 1 The advanced, intermediate and beginning students perform differently on three types of listening-speaking test tasks at the .05 level.

Hypothesis 2 The three task types of the CMLST and groups of students differently affect the oral abilities of the three groups of students at the .05 level.

Hypothesis 3 The three groups of the test takers perform on each type of CMLST differently in terms of accuracy, fluency, complexity, and comprehension.

1.5 Scope of the Study

Population : The population of this study is 253 first year students of BRU enrolling in Communicative English I course in the first semester of the academic year 2009. They study in the Faculties of Education as well as Humanities and Social Sciences from 2 majors: English and Business English.

The instruments : The Computer - mediated Listening - Speaking Test Tasks (CMLSTT) is a semi-direct computer-mediated listening - speaking test task. The students do not interact directly with the interlocutor but listen to the prompts and the questions from the video clips delivered by the computer. Test takers talk into the microphone via the computer. The content of the test is taken from the course description and objectives of "Communicative English I" course .

The task types of the speaking test : The three task types are chosen from the language activities mostly used in "Communicative English I" course. The language functions of each task type are developed from the course description and objectives of this course.

1.6 Assumptions of the Study

The students who participate in this study have some basic knowledge of using computers, and have the same language background with no experiences abroad and no exposure to English outside classes. They have a similar number of years in learning English and similar learning habits in practicing English outside classes that do not enhance their oral abilities.

1.7 Definition of Terms

The CMLSTT is a computer - mediated listening - speaking test that contains three task types. The video clips of three different situations of three language functions are presented to the test takers via the computer. The test takers watch the video clips, and listen to the instructions from the computer. They assume the role as an interlocutor in the video clips and answer the questions by recording their responses into the microphones. Their responses are assessed by two raters in terms of four aspects of language performance: accuracy, fluency, complexity, and comprehension.

KET is Key English Test of Cambridge ESOL. It is an exam for people who can use everyday written and spoken English at a basic level. It covers all four language skills — reading, writing, listening and speaking. Only listening and speaking papers are selected in order to classify the students according to their oral abilities. The listening part lasts 30 minutes and the speaking part takes approximately 10 minutes. The mark of each part counts 25% of the total test.

The task type of CMLSTT refers to three language functions: answering questions, describing pictures, and transferring information which are derived from language activities in "Communicative English I" course. The language functions are taken from the course description and objectives of this course. The contents and functions of each task can be assumed to be familiar to all subjects.

Answering questions task refers to the task including 5 open-ended questions about the test takers' personal information asked by the interviewer in the video clip. Each question is heard twice. The test takers have to answer all of the questions by using their own information after they hear all of the questions and within the given time.

Describing pictures task refers to the task that the test takers study the pictures of personal information of a Thai male athlete. The test takers look over the pictures and use them as the guidelines to make sentences describing his personal information.

Transferring information task refers to the task that the test takers have to read the guided words in the given table about the personal information of a Thai female athlete and they create their own sentences about her from those words.

Oral abilities refer to the CMLSTT scores obtained from evaluating four aspects of language performance : accuracy, fluency, complexity, and comprehension. The students can comprehend the instructions and questions in English and react to the questions in order to achieve the communicative goals in different language functions.

Accuracy means the CMLSTT scores assessed from the students' ability in showing the degree of accurate grammar, appropriate vocabulary and appropriate pronunciation to convey the intended meaning.

Fluency refers to the CMLSTT scores assessed from the students' ability in speaking fluently and intelligibly to produce smooth and natural flow of speech.

Complexity is the CMLSTT score assessed from the students' ability in giving detailed responses with a wide range of connectors and cohesive devices.

Comprehension is the CMLSTT scores assessed from the students' ability in reacting to instructions and questions by giving the adequate and relevant responses in order to achieve the communicative goals of each task.

Groups of students refer to the students classified according to their KET scores into three groups, namely, advanced, intermediate, and beginning groups.

The advanced students are the students whose KET scores are at or above the level of 80th percentile or the scores above or equal to 26 points.

The intermediate students are the students whose KET scores are between 70th and 40th percentile or the scores between 16-22 points.

The beginning students are the students whose KET scores are at or under the 25th percentile or the scores below or equal to 14 points .

1.8 Significance of the study

Regarding theoretical contribution, the findings of this study will confirm the effects of task types of the Computer-mediated Listening-Speaking Test Task (CMLSTT) on English oral abilities. This will yield more information on the constructs and design of CMLST to suit the test takers with diverse oral abilities and the four aspects of language performance to be assessed. In terms of practical contributions, the CMLSTT developed in this study is expected to be a useful tool for "Communicative English I" course in assessing the oral abilities of a large number of the first-year English major students at Buriram Rajabhat University via computers. Furthermore, the test data can help develop new ways of teaching, material design, and assessment of listening and speaking courses to serve Rajabhat University students with different oral abilities.

Chapter II

LITERATURE REVIEW

This chapter presents a review of the related literature concerned with the underlying concepts of this study. It covers reviews of the components of speaking skills, the theoretical framework for assessing oral abilities, types of oral assessment and how to score oral abilities, task-based assessment for oral abilities and computer-assisted language assessment.

2.1 The Components of Speaking

Among the four language skills (listening, speaking, reading and writing), speaking is paid less attention than other skills since it is commonly assumed that everyone can do spontaneously with "colloquial register" which is a level of formality we use in talking to people in our daily life or informal language. In addition, it is rather difficult to study the spoken word due to its nature to be transient, improvised and dynamic (Bygate, 1987: vii). People can use speech to convey what they would like to converse such as to argue, to persuade and to convince through their body languages, eye contact or intonation (Hughes, 2002). Thus, speaking is a skill we should take into account to support learners to be able to communicate in their basic transaction by speaking confidently (Bygate, 1987). Speaking in many educators' views refers to the basic medium of communication which has a great influence on language use. Florez (1999 cited in Bailey and Nunan, 2005), for example, defined speaking as the process of producing, receiving and processing information in order to convey the meaning. Relevant to the concept of testing, according to Fulcher (2003: 23), speaking is "the verbal use of language to communicate with others that can be observed and measured or scored".

Most language teaching courses at the beginning to intermediate levels put more emphasis on teaching conversation or dialogues though there are other crucial forms of spoken English apart from the English conversation. Brown (2001) proposes two forms of oral language adapted from that of Nunan (2001): monologue and dialogue. Monologue occurs when one speaker gives speech at any length such as in opening speeches, lectures and news reports. The speech has continued without interruption or paying attention to the listeners' comprehension. The problems in

comprehension result from unplanned monologues other than planned ones. There are two or more speakers in dialogues which are composed of two main purposes i.e. interpersonal and transactional dialogues. “Interpersonal or interactional dialogues” (Brown, 2001: 251) express social relationships between the speakers and the listeners while the speakers intend to exchange the factual or propositional information in “transactional dialogues”. Unfamiliarity between participants plays an important role in the effectiveness of their communications. Thus, they have to produce conversations more explicitly. In actual conversations teachers cannot predict what kinds of conversations learners have to deal with or whether the interlocutors are strangers or acquaintances. For this reason, in any speaking courses teachers should include a variety of spoken discourse in order to make sure that students can converse in many situations.

Apart from two forms of oral language which teachers have to take into consideration when designing speaking courses, microskills and macroskills of oral communication are other concerns when developing speaking lessons. Brown (2004) presents eleven microskills and five macroskills shown in Table 2.1.

Table 2.1

Microskills and Macroskills of Oral Communication

(Brown, 2004: 142-143)

Microskills
1. Produce differences among English phonemes and allophonic variants.
2. Produce chunks of language of different lengths.
3. Produce English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours.
4. Produce reduced forms of words and phrases.
5. Use an adequate number of lexical units (words) to accomplish pragmatic purposes.
6. Produce fluent speech at different rates of delivery.
7. Monitor one’s own oral production and use various strategic devices---pauses, fillers, self-corrections, backtracking---to enhance the clarity of the message.

Table 2.1 (cont.)

Microskills and Macroskills of Oral Communication

(Brown, 2004: 142-143)

8. Use grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement, pluralization), word order, patterns, rules, and elliptical forms.
9. Produce speech in natural constituents: in appropriate phrases, pause groups, breath groups, and sentence constituents.
10. Express a particular meaning in different grammatical forms.
11. Use cohesive devices in spoken discourse.

Macroskills

12. Appropriately accomplish communicative functions according to situations, participants and goals.
13. Use appropriate styles, registers, implicature, redundancies, pragmatic conventions, conversation rules, floor-keeping and-yielding, interrupting, and other sociolinguistic features in face-to- face conversations.
14. Convey links and connections between events and communicate such relations as focal and peripheral ideas, events and feelings, new information and given information, generalization and exemplification.
15. Convey facial features, kinesics, body language, and other nonverbal cues along with verbal language.
16. Develop and use a battery of speaking strategies, such as emphasizing key words, rephrasing, providing a context for interpreting the meaning of words, appealing for help, and accurately assessing how well your interlocutor is understanding you.

Microskills include the abilities in producing smaller language elements such as phonemes, words, etc. with the belief that these pieces can make up the whole pictures. Being more complex than microskills, macroskills refer to larger parts of language, namely, discourse, strategic devices, sociolinguistic features and fluency.

Their difficulties are based on the stage and context of the test takers. Micro and macroskills can be used not only in teaching speaking but in speaking assessment. It can be used as the objectives of the test or a checklist of the objectives in evaluating oral abilities.

Writing a good speaking test is not easy and it has not been done much due to the fact that speaking is a difficult language skill to be assessed reliably since it depends on different perspectives related to the factors that show how well students can speak English. It also relies on the personal variables such as proficiency levels (Fulcher, 2003; Luoma, 2004). In addition, there is little pedagogical research supporting the effectiveness of specific methods in teaching oral abilities or any best way to test oral abilities. In order to find the solution, test designers have to set the components of oral abilities or "construct definition". Test purposes together with testees' needs and motivation will be used in designing the speaking construct, content and test inferences. Moreover, the construct should serve the score users' needs. As the speaking process is so complicated, we cannot design the complete speaking test that covers all of the components of oral ability. The test construction only relies on the purpose of the test and the language ability to which we infer the scores (Fulcher, 2003). The construct of the test depends on the theoretical models on which we choose to base the test and this is discussed as follows.

2.2 Theoretical Frameworks for Assessing Oral Abilities

For assessing oral abilities, it is worth exploring some relevant theoretical frameworks which include communicative competence and its limitations (Hymes, 1971, 1972), communicative language ability (Canale, 1983; Savignon, 1997) and communicative model of language ability including Bachman and Palmer's (1996).

2.2.1 Communicative Competence

Based on the language acquisition research during the 1970s and 1980s, both native speakers and second language learners do not make sentences by putting together the language components they learn, namely nouns, adjectives and verbs. They rather learn the language elements from conversing with people (Bailey and Nunan, 2005). That leads them to give and take the information among themselves or what we call "communication". Canale (1983: 4) views communication as "the exchange and negotiation of information between at least two individuals through the use of verbal and non-verbal symbols, oral and written-visual modes as well as

production and comprehension processes". Savignon (1997: 14) states communication as "a continuous process of expression, interpretation and negotiation of meaning". Communication covers all of the modes that interlocutors use in conversing, interpreting information and finding an agreement leading to the communicative goals. To achieve the goal of communication, both language producers and receptors are needed. In order to interpret the test scores meaningfully and get broader information about test takers' language ability, the test should be designed depending on theoretical frameworks to which the examiners will relate learners' language performance. Moreover, those theoretical frameworks should not put an emphasis only on the micro skills which require knowledge of basic language elements, namely sounds, vocabulary and grammar since it is not enough for learners to effectively communicate with others in English. Among the numerous theoretical models, "communicative competence" developed by Hymes (1971, 1972) is the one that is very popular among the test designers because it involves understanding of language components and how to appropriately use language in the communicative context.

To decide whether language users have "communicative competence", Hymes (1971, 1972) proposes four questions in the analysis of language uses based on "The Theory of Language Users and Language Use" as follows :

1. Is the use of language "possible" in grammatical, cultural and communicative way? For example, the language users know that "the was cheese green" is grammatically incorrect.

2. Are the sentences " feasible " within language situations and ways to communications? Some expressions are grammatically correct but are not appropriate in their socio-cultural contexts or it is not common for the people in that community such as greeting people with "good bye" or closing the TV program with "may god be with you" instead of good-bye or bye-bye.

3. Is the language used " appropriate " to the speech events and language contexts?

4. Are the languages actually used in each society? Some expressions are possible and appropriate but they are not accepted in certain social groups.

Although the sentences are possible, feasible and appropriate, they will be meaningless if those sentences are not typically used by the people in that society (the question no.4 is neglected). Despite the fact that this theory is accepted

widely in the field of language learning, its details have not been implemented much due to the abstraction of the idea. With the belief that the success in oral communication does not depend on only a single factor but rather the numerous components, Littlewood (1981) proposes the following four main domains in the communicative skills of language learners:

1. The learner has " Linguistic Competence " at the level of the capability of communicating what he would like to say.
2. The learner can understand both linguistic and communicative functions of language forms. For example, the question "Why don't you close the door?" has the meaning beyond the reasons of not to close the door. The function can be a command, a suggestion or a complaint.
3. The learner can develop strategy skills in real situations.
4. The learner must comprehend the social meaning of language forms in order to produce language appropriately according to their real uses.

Communicative language ability according to Littlewood's (1981) and Hymes' views (1971, 1972) refers not only to the ability in deciding which sentences are literally correct and the ability in selecting the sentences really used in the specific society but also to the ability in producing the languages responding to their language functions. Thus, the communicative competence includes "functional" and "structural" aspects of language. Since the sentence structure is stable and straightforward while the function of language varies with the social language environment, it is rather difficult to prepare the language functions in advance. Some linguists extend the scope of communicative competence to the ability to use the language appropriately in the real language situation. Canale (1983), for instance, categorizes the components of communicative ability or competence into two broad parts: knowledge and skill. The former refers to what one knows about language and other aspects of communicative language use. And the latter is how well one can perform this knowledge in actual communication. Similarly, Savignon (1997: 15) classifies communicative ability or competence into two main components with the same concepts but different names; that is, "what one knows" or "a presumed underlying ability" is called "competence" and "performance" for "what one does " or " the overt manifestation of that ability". As a result, in order to decide whether learners have communicative ability or competence, it is based on the knowledge of language together with the knowledge of how to use it properly in the real situations.

In addition to competence and performance, Savignon (1997) also adds the important characteristic of the communicative competence, the “interpersonal traits” of communicators with the same symbolic systems which can be applied to all language skills. Furthermore, it is the language interpretation based on different contexts together with the learners’ background knowledge.

Canale (1983) categorizes knowledge and skill into 4 groups as follows:

1. Grammatical Competence involves features or rules of language in order to comprehend and produce the utterances literally. To clarify this, Savignon (1997: 41) explained that grammatical competence is “the mastery of the linguistic code, the ability to recognize the lexical, morphological, syntactic and phonological features of language and to manipulate these features to form words and sentences”. Whether or not the sentence is grammatically correct depends on whether these structures are in common among native language users. As a result, grammatical competence is expressed not by stating a rule of grammar but using it.

2. Sociolinguistic Competence is concerned with the way to understand and express the utterances appropriately according to “the contextual factors”, e.g. the purposes and the content of the conversation, functions of interactions and participants’ roles. To illustrate this, Savignon (1997) stated that this knowledge includes the ability in the judgement of appropriateness of language use and involves not only what and how to say but when to be silent or incompetent as well.

3. Discourse Competence refers to the ability to use the knowledge of grammatical forms as well as meanings to unify different types of texts via coherent (how to link the meanings of the utterances in a text) and cohesion (how to comprehend the texts by relating them grammatically). This knowledge is useful in comprehending different genres such as a telephone conversation, television commercials or office memos.

4. Strategic Competence includes both verbal and non-verbal communication used as the compensatory or survival strategies for the breakdown of interactions resulting from limited factors such as distraction, inattention in communication or imperfect knowledge of rules and the strategies for more successful interactions.

Even though Canale’s idea is well-accepted among linguists, it was criticized by some theorists such as Skehan (1998) in that the four components of competences cannot be related to broader language performance or different contexts. Moreover,

the strategic competence is rather limited in being the compensatory strategies only in problematic circumstances but not in ordinary communication.

2.2.2 Communicative model of language ability in language testing

There are some problems in implementing the theories of communicative competence mentioned previously due to the abstraction of the ideas, the limited language ability and contexts to which the examiners relate the scores. “The communicative model” or “the model of language ability” in language testing presented by Bachman and Palmer (1996) has been mostly used in language learning and assessment. The model of language ability consists of four components of language use: 1) personal characteristics or test takers’ characteristics which influence their language performances; 2) topical knowledge or real world knowledge which helps facilitate learners to use language in different situations; 3) affective schemata or emotional factors that affect the way examinees react to the test task flexibly or limitedly and 4) the ability that helps learners make and comprehend sentences or language ability including two components: language knowledge and strategic competence. Language knowledge is used by metacognitive strategies in making and comprehending utterances while strategic competence is the language users’ metacognitive organization and the monitoring of the situation. Figure 2.1 shows the process of language use and language test performance with the four main individual characteristics of language users included in the inner circle. They interact with one another and are linked by strategic competence. Also, it interacts with the characteristics of language use or testing situation. Among all individual characteristics, language ability including language knowledge and strategic competence is the most salient one which the test examiners always infer from the test performance.

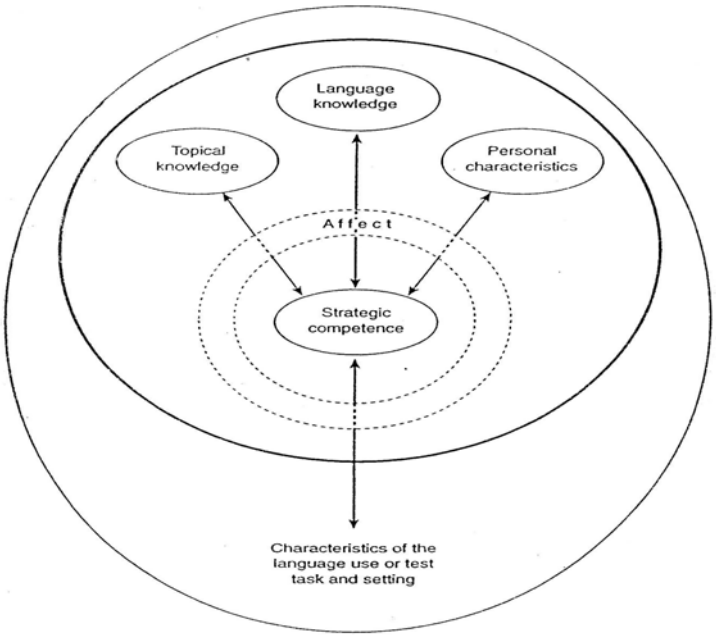


Figure 2.1: The interaction of the components of language use with language test performance (Bachman & Palmer, 1996: 63)

Language knowledge presented in Table 2.2 (Bachman and Palmer, 1996: 68) includes two main categories: organizational and pragmatic knowledge. The former puts an emphasis on the knowledge of organizing the sentences or utterances while the latter gives weight to the relationship between language forms language users use to achieve their goals and language situations.

Table 2.2

Area of Language Knowledge

(Bachman and Palmer, 1996:68)

Organizational knowledge (how utterances or sentences and texts are organized)
Grammatical knowledge (how individual utterances or sentences are organized)
Knowledge of vocabulary Knowledge of syntax Knowledge of phonology/graphology
Textual knowledge (how utterances or sentences are organized to form texts)
Knowledge of cohesion Knowledge of rhetorical or conversational organization

Table 2.2(cont.)

Area of Language Knowledge

(Bachman and Palmer, 1996:68)

Pragmatic knowledge
(how utterances or sentences and texts are related to the communicative goals of the language user and to the features of the language use setting)

Functional Knowledge
(how utterances or sentences and texts are related to the communicative goals of the language user)

Knowledge of ideational functions
Knowledge of manipulative functions
Knowledge of heuristic functions
Knowledge of imaginative functions

Sociolinguistic knowledge
(how utterances or sentences and texts are related to features of the language use setting)

Knowledge of dialects/ varieties
Knowledge of registers
Knowledge of natural or idiomatic expressions
Knowledge of cultural references and figures of speech

Organizational knowledge is classified into two areas: grammatical knowledge which includes knowledge of vocabulary, syntax, phonology and graphology in producing formal utterances while the knowledge in producing text or textual knowledge involves the knowledge of relationships among sentences(cohesion) and the organizational text development (rhetorical or conversational organization).

Pragmatic knowledge includes functional knowledge which is the ability in interpreting the relationships between sentences and the users' purposes and sociolinguistic knowledge which interprets language use suitable for language situations. The functional knowledge comprises four functions of language use (ideational, manipulative, instrumental, and imaginative). Utterances that perform an ideational function express language users' experiences of the real world, that is to say the ideas, knowledge or feelings are exchanged. The language used to affect the world around them has a manipulative function. And language used to extend the knowledge of the world around them expresses a heuristic function. An imaginative function refers to the language used to create the world for aesthetic or humorous purposes. Sociolinguistic knowledge composes the knowledge of dialects and varieties, registers, natural or idiomatic expressions, and cultural references and figures of speech.

Pragmatic competence or Pragmatic knowledge according to the Council of Europe (CEF, 2001) does not include sociolinguistic competence as Bachman and Palmer (1996) do, but includes “discourse competence”, the ability in arranging the sentences in the order to create the coherent texts, and “design competence”, the arrangement of the text resting on the patterns of social interactions (interaction schemata). The main component of pragmatic knowledge that affects the language users in communicating for specific functions is “Functional Competence”.

In actual conversation, language users need the exchanges of single sentences to serve specific purposes (microfunctions). The Council of Europe (2001: 126) categorizes the microfunctions into six groups as follows :

1. giving and asking for information, e.g. identifying, correcting, reporting
2. expressing and asking the point of view, e.g. pleasure /displeasure, sympathy, preference, ability, probability, agreement/disagreement
3. suasion, e.g. warning, encouragement, invitation
4. socializing, e.g. attracting attention, introduction, toasting
5. structuring discourse, e.g. opening, turn taking, closing
6. communication repair, e.g. appealing for assistance, paraphrasing, repetition

In order to continue the conversation to the end, only using language as microfunctions is not enough; they also need to interact with their interlocutors with the spoken discourse arranged to serve the same purposes (macrofunctions) such as description, demonstration, argumentation. It is believed that the more microfunctions the learners can gain control of, the more they learn the suitable way to talk in different situations. Since the language functions affect the speaking task design, they are used as the framework for formal speaking tests like Test of Spoken English (TSE)(ETS, 2001). The language functions in TSE are closely related to those of microfunctions by the CEF (2001). The test also includes different situations and relates the language functions to those that the test would like to refer to.

Not only the language functions used in designing speaking tests, but the knowledge of using the correct patterns of communication or “interaction schemata”, e.g. question-answer, request-acceptance/ non-acceptance test designers should also take into consideration. In other words, in order to test speaking ability learners need to demonstrate the abilities to use language to do “interaction skills” which compose two kinds of subskills: routines and improvisation skills (Bygate, 1987). The way in

which speakers organize their speech in typical patterns or “conventional ways of presenting information” are routine skills. Two main kinds of routines, namely, information and interaction routines proposed by Bygate (1987) are counted as the salient qualification the students need for communicating with both strangers and the familiar faces (Brown, 2001). Information routine refers to the structures that are often used in presenting the ideas e.g., stories, description of places and people, presentation of facts, comparison, instructions. This structure can be classified into expository routines and evaluative routines. Expository routines concern with the arrangement of the factual information or the identity of the subject such as narration, description and instruction whereas the principal types of evaluative routines or making reasonable conclusions are explanations, predictions, justifications, preferences and decisions. Instead of depending on the content as information routines, “interaction routines” put more emphasis on the sequences of factors happening in different situations. Thus, they include the interactions such as telephone conversations, interviews which tend to be arranged according to their characteristics.

When language speakers encounter problems in communication, in cases where only routine skills are not be enough to make them understood, improvisation skills will also be used. Learners need to develop skills in the negation of meaning and management of interaction which are two main kinds of improvisation skills .The skill speakers and listeners use to follow the procedures to check their understanding is the skills required in negotiation of meaning. The management of interaction includes the interlocutors’ rights to differentiate the kinds of speech e.g., a lecture, a conversation or meeting and to decide when and how to take the floor, when to introduce a topic or change the subject, how to keep a conversation going as well as when and how to end the conversation.

In spite of the fact that Bygate’s (1987) idea of interactive speech is not a comprehensive and accessible working model of language use, it has provided language activities and interactions which were an appropriate framework at that time (Weir, 2005). Following Luoma (2004), “the communicative model” that has been well-accepted in language learning and assessment is the one proposed by Bachman and Palmer (1996). In order to develop a speaking test based on applied purposes, speaking assessments can hardly cover every aspect of a theoretical model. An eclectic approach, of course, is the natural way for test developers to implement the model to the test. For the Computer-mediated Listening-Speaking Test Tasks

(CMLSTT) used as a main instrument in this study, only organizational knowledge and the knowledge of ideational functions, a component of functional knowledge were selected to be the constructs of the test. Due to the fact that the test is based on the course description of Communicative English I Course which is related only to the simple language forms and functions, sociolinguistic knowledge is not included in the test constructs. Regarding language functions, the framework of CMLSTT was derived from the contents of the course. Since the length of the test should not last so long in order not to bore the test takers, only the language function (giving personal information and talking about daily routine) were selected from the responses of the questionnaires given to English teachers at Buriram Rajabhat University.

2.3 Types of Speaking Tests

In designing speaking tests, it is important to explore tremendous factors which affect the test takers' speaking performances. One of the crucial factors that should be considered is "speaking test types" which include how to choose the types of questions appropriate to elicit the expected answers, the proper modes of testing as well as the way in which the test items are written to serve the test objectives. The speaking test types mentioned in the literature are discussed below.

2.3.1 Direct Tests

The direct test was first introduced by Clark (1975 cited in Shohamy, 1994: 100) as "test format and procedures which duplicate the setting and operation of real-life situations in which the proficiency is normally demonstrated". This kind of test setting has to include the examinee and one or more human interlocutors communicating in oral interactions (Shohamy, 1994). Based on Bailey and Nunan (2005), a direct test refers to the test which lets the students contact face to face to the interlocutors such as interview, conversation, unscripted role-play whereas Luoma (2004) cited the same concept of this kind of test with a different name, "live, face to face interaction". The interaction occurring in this kind of testing is bidirectional, i.e. each speaker's turn reacts to the previous turn and clarifies the utterances. "Spoken interaction"; therefore, is the main construction of this test. A popular example of direct speaking tests is the Oral Proficiency Interview (OPI) developed by the American Council on the Teaching of Foreign Languages (ACTFL).

2.3.2 Indirect Tests

In this kind of tests, students do not speak but are tested via other language skills tests. In a conversation cloze test, for example, students' speaking ability is assessed via the written text. It is assumed that students who can choose the proper words for the conversation context will have a high ability in speaking. Other examples are paper-based speaking or phoneme discrimination tasks (Bailey and Nunan, 2005).

2.3.3 Semi-direct Tests

Semi-direct tests were developed to control the extraneous variables occurring during conducting direct tests such as the role relationship, interviewees' and interviewers' personality and gender, the purpose of the interaction. In this kind of test, students do not need to interact directly with the interlocutors; in fact, they respond to the prompt from the tape-recorder (Bailey and Nunan, 2005). Unlike the direct test, "tape-based testing" cited by Luoma (2004) is the one-way method which accommodates the examinees to the tape. The construction of the test is concerned with "spoken production". The Semi - direct Oral Proficiency Interview Test (SOPI) is an example of such a semi-direct test which stimulates the test takers by using recorded and visual tasks to respond orally.

Regardless of the problems resulting from the unreliability in interlocutor performances, direct testing is accepted as the most valid method of assessing speaking. But in some testing situations, for example, in an entrance examination where there are a lot of test takers attaining the tests or a few proctors have to administer the tests in different locations, then direct testing is impractical. On the other hand, semi-direct tests play crucial roles to compensate in these situations. Though the Semi-direct Oral Proficiency Interview test (SOPI) is used as an alternative for the direct test, the Oral Proficiency Interview (OPI) in some situations, it is questionable whether the two tests measure similar things or whether they are correlated. Shohamy (1994) studied the validity of OPI and SOPI in quantitative and qualitative ways. Though the high correlation between these two tests was found, they are different in some aspects after analyzing the data qualitatively. SOPI included the main language functions, namely discussing advantages/disadvantages, expressing opinions, apologizing and so forth while OPI elicited the functions like supporting opinions, elaborating, expressing courtesy, complaining, reporting and so on. In terms of language functions and topics elicited by the two tests, low-proficiency test

takers attaining SOPI had an advantage of exposing to language functions and topics over those doing OPI. The high-level test-takers from the OPI had an advantage to use a wider range of language functions and topics. Therefore, direct and semi-direct tests are not interchangeable since SOPI and OPI are comparable in some areas while in others they are not (O′ Loughlin, 2001). Shohamy (1994) did not include which one is better or more valid. In order to design valid assessment, both direct and semi-direct tests are required. Moreover, the appropriateness depends on the test purposes and intended use of the test and the test results. As O′Loughlin (2001) suggested that the semi-direct test is suitable for more specialized testing contexts such as the assessment of monologic speaking ability.

Luoma (2004) also adds some kinds of task types based on to what extent the instructions provide the guidelines for the test takers. Following are some task types.

1. Open-ended speaking tasks

The main objective of this kind of testing is allowing the students to perform language in activities such as giving a presentation or making a request. The Oral Proficiency Interview (OPI) is an example of "description tasks". The open ended speaking tasks can be either direct or semi - direct depending on whether there is the relationship between the examinee and other interlocutors or not.

2. Semi-structured tasks

They refer to the "reacting in situations". In a role play activity, for example, the test taker has to assume himself/ herself to be a part in the language situation and is asked to respond to the given prompt. Students have to use both language knowledge and language use. Thus, a semi-structured task is more suitable for tape-based testing than face-to face testing due to numerous language situations.

3. Structured tasks

The answer from this kind of testing is short and fixed under teachers' expectations. Despite the lack of creative answers and freedom in answering, students' scores can be compared and it is easy to check the answers by using a scoring key. The examples of this kind of testing are reading aloud, sentence repetition such as in the Phone Pass. Another interesting example called "reacting to phrases" is concerned with the typical question-answer or comment-response sequences such as "adjacency pairs", e.g. refuse-accept the offer, give-reply thanks. The test takers have to respond to the utterances from the tape in order to continue the

social activities properly. The structured tasks are usually assessed by using tape-based testing.

Apart from the test types distinguished by the test modes described above, Weir (2005) uses the same terms: direct and indirect tests to refer to the relationship of the test content compared with language ability in real-life situations. The test that has similar relationship between the content as well as construct of the test and real life performance is called a direct test and the test performance can be compared to the target performance. Information gap exercises, free or controlled interviews and monologic tasks are good examples of this kind of test. In contrast, an indirect test will cause difficulty in score reference to language performance in real-life situations. The examples of this testing are information transfer and mini-situations on tape. In Table 2.3, Weir (2005) also presents the advantages and disadvantages of these test formats the test administrators should take into account before deciding the most suitable ones for their students and learning contexts.

.

Table 2.3
Advantages and Disadvantages of Test Formats
(Weir, 1993; 2005)

Test Formats	Characteristics	Advantages	Disadvantages
1. Indirect Test			
1.1 Information Transfer	The test-takers describe a chart or a figure	- It is straightforward, is preferred by School Exam Board in Britain and accepted as one of the most effective test types. - This test type can avoid the interference in measurement of other language skills.	- It is difficult for teachers to prepare the suitable materials for this kind of test. - The students can express only informational routine skill which can be used in the limited situations e.g. oral presentation or lecture.
1.2 Mini-situation on Tape	The test-takers respond to a number of remarks made to them or the situations they might find themselves in.	- The test takers experience various speech events (e.g. different interlocutors, accents, settings, roles, topics). - The scoring is more reliable as the responses can be replayed. - It's practical in marking at any convenient time and place.	- Due to the flexible language situations, it might be unfair for the students.

Table 2.3 (Cont.)
Advantages and Disadvantages of Test Formats
(Weir, 1993; 2005)

Test Formats	Characteristics	Advantages	Disadvantages
1.3 Narrative on a picture sequence	The test takers describe a series of pictures in sequences.	- Students can give the extended speech useful for assessing many aspects of criteria particularly the ability in grammar. - Students can give informational routine.	- The candidates' potential may be affected by ineffective materials e.g. unclear pictures or those influenced by cultural or educational bias. - It is not authentic as students rarely meet this kind of situation in their real life. - Students cannot express their improvisational skills or abilities in interaction.
2. <u>Direct Test</u> **<u>The Interaction between Students and Students</u>**			
2.1 Information Gap Exercise	Students work in pairs and each has to be responsible for his/her part to get the missing information.	- The activity includes informational, interactional routines and improvisational skills.	- It may cause some problems in scoring if one interlocutor takes a longer turn than the other.

Table 2.3 (Cont.)
Advantages and Disadvantages of Test Formats
(Weir, 1993; 2005)

Test Formats	Characteristics	Advantages	Disadvantages
The Interaction between Students and Examiners or Interlocutors **2.2 Free interview/ conversation**	There are no structures or sets of procedures the conversation has to follow.	- Students can create their own utterances similar to those in real life with purposeful and unpredictable interaction. - It is the most appropriate classroom test for responsible students. - It prevents the problem of students' lacking topical knowledge. - The test writers can include many test theories and contexts and easily make in various forms depending on numerous types of information - The stimuli, namely, paces, scopes and levels of interaction are flexible.	- The candidate may dominate the conversation if he/she is more familiar with the topic. - It is impractical in test administration and the maintenance of test security. - The test cannot include the wide range of situations due to the limited variables of interviewer.

Table 2.3 (Cont.)
Advantages and Disadvantages of Test Formats
(Weir, 1993; 2005)

Test Formats	Characteristics	Advantages	Disadvantages
2.2 Free interview/ conversation(cont.)		- The test-takers can create their own topics and directions of the interaction.	- The reliability problem may occur because of the candidates' flexible answers. - Due to the practicality of this method, a large number of students take this test,so it takes more time in test administration. - It is not suitable for some students who are shy and introvert. - There are some problems in scoring the discourse construct. - This test is less authentic as the interviewer tries to ask questions that elicit the interviewees' answers.

Table 2.3 (Cont.)
Advantages and Disadvantages of Test Formats
(Weir, 1993; 2005)

Test Formats	Characteristics	Advantages	Disadvantages
2.3 The controlled interview	The interviewer dominates the conversation in selecting questions, developing topics and allowing the candidates to respond to the questions.	- The test is more reliable than the free one as the candidates are asked the same questions and their answers can be compared.	- The interviewee has no chance to initiate the questions and topics, manage the conversation and keep the conversation going. - It takes more time and money in test administration with a large number of candidates.
2.4 Monologic tasks	The candidates will be stimulated by the questions or situations and they have to give the extended utterances.	- It can be used to measure more language components because of the longer utterances. - The test activities are related to real language situations due to the authentic contexts. - Regardless of the co-construction of discourse, this kind of test can be used to study the influence of "intra-task variation"(e.g. task demand, setting)	- It takes more time and money in test administration with a large number of candidates.

With regards to many types of speaking tests, test developers can choose which test can best serve the objectives of their speaking assessment. They can choose only one category or they can combine all in one test depending on what they want speaking to be like. Each type has its unique benefits. It depends on what you want to give more weight in your test. If you would like to emphasize the interactions among the examinees, the direct test will be essential for this. But if the examinees are not fluent in speaking, it is better to start with a structured test. For the first year students of Buriram Rajabhat University who are at the beginning level of English, the semi-structured tasks can better assess their language performance. Though they cannot reach the level in using language freely as in the open-ended test, a semi-structured test can pave the way for them to the higher level. Numerous language situations in this kind of testing seem to be the preparation stage for them before facing the real world. Therefore, CMLSTT was designed for this purpose. Due to the problem of few test administrators with a large number of students, the test was based on the semi-direct test which could let the test takers interact with the computer instead of the interviewers. The three test formats (answering questions, describing pictures and transferring information) were selected according to the language activities mostly used in class. As a kind of a controlled interview, answering question activity was chosen to make the test reliable so the candidates were asked the same questions. In terms of describing pictures, this format is hoped to be the first stage to stimulate the students to speak though it is not actually authentic. Regarding transferring information, it can help control extraneous variables from other language skills.

2.4 Speaking Rating Scales

After the students take the speaking test, the scores obtained are interpreted according to speaking rating scales which are used to tell how well the test takers can speak. Rating scales are statements at any length which can explain the meaning of the scores written along the continumn (Luoma, 2004). Rating scales composed of many levels describing the test takers' language performance are viewed as the operational definitions of linguistic constructs. It is specific to the purpose of the test and group of the examinees. Therefore, people who would like to use the scales should be trained (Davies et al., 1999).

It is not easy to write speaking rating scales with reliable evidence about language and it is rather difficult to make the short form of rating scales that are

practical to all users. In addition, there are a few common scales everyone can use. Due to the fact that tests vary in purposes as well as the target groups, tests should be designed in various versions serving the purposes and needs of the users. Rating scales; therefore, are written to respond to the three groups of users: rater-oriented scales, examinee - oriented scales and administrator - oriented scales (Luoma, 2004; Alderson, 1991; Fulcher, 2003).

1. **Rater or assessor - oriented scales** : These scales guide the raters to decide consistently based on the expected performance . They should include the construct definition used in the rating process in the limited time such as in a face-to-face speaking test

2. **Examinee or user-oriented scales** : These scales show the information about the overall behavior of a test taker and his specific strengths and weaknesses at a given level. The descriptors propose “what the learners can do in the second language” or “can-do statements”.

3. **Administrator or constructor-oriented scales** : These scales give detailed information of the task along with the task types used in eliciting meaningful language samples. This information is used in selecting the task to be included in the test.

Weir (2005) and Luoma (2004) present a good example of rating scales prepared for the multiple user groups which is the Test of Spoken English (TSE) by the Educational Testing Service (ETS, 2002) presented in Table 2.4.

Table 2.4

Test of Spoken English Rating Scale

(ETS, 2002: 30)

60 Communication almost always effective: task performed very competently.
Functions performed clearly and effectively
Appropriate response to audience/ situation
Coherent, with effective use of cohesive devices
Use of linguistic features almost always effective;
communication not affected by minor errors

50 Communication generally effective: task performed competently.
Functions generally performed clearly and effectively
Generally appropriate response to audience/ situation
Coherent, with some effective use of cohesive devices
Use of linguistic features generally effective;
communication generally not affected by errors

40 Communication somewhat effective: task performed somewhat competently.
Functions performed somewhat clearly and effectively
Somewhat appropriate response to audience/ situation
Somewhat coherent, with some use of cohesive devices
Use of linguistic features somewhat effective;
communication sometimes affected by errors

30 Communication generally not effective: task generally performed poorly.
Functions generally performed unclearly and ineffectively
Generally inappropriate response to audience/ situation
Generally incoherent, with little use of cohesive devices
Use of linguistic features generally poor;
communication often impeded by major errors

20 No effective communication: no evidence of ability to perform task
No evidence that functions were performed
No evidence of ability to respond to audience/ situation
Incoherent, with no use of cohesive devices
Use of linguistic features poor;
communication ineffective due to minor errors

The communication abilities in the first line in Table 2.4 are the versions for the administrators. And four additional statements below each of the bold statements are the examinees' communicative competence: functional, sociolinguistic, discourse and linguistic competence and this is the version for the examinees and the public. Moreover, ETS provides the new raters more levels of details with the information of the type of language as presented in Table 2.5.

In summary, differences between versions involve the terminology, the details of descriptors and examinees' language performances. The scale emphasizes the levels of these abilities (Council of Europe, 2001).

Table 2.5

Test of Spoken English Rating Scale for Overall Features

(ETS, 2002: 31)

60	**Communication almost always effective: task performed very competently.**	Speaker volunteers information freely with little or no effort, and may go beyond the task by using additional appropriate functions. * Native-like repair strategies * Sophisticated expressions * Very strong content * Almost no listener effort required
50	**Communication generally effective: task performed competently.**	Speaker volunteers information sometimes with effort, usually does not run out of time. * Linguistic weaknesses may necessitate some repair strategies that may be slightly distracting * Expressions sometimes awkward * Generally strong content * Little listener effort required
40	**Communication somewhat effective: task performed somewhat competent.**	Speaker responds with effort; sometimes provides limited speech sample and sometimes run out of time. * Sometimes excessive, distracting, and ineffective repair strategies used to compensate for linguistic weaknesses (e.g., vocabulary and/or grammar) * Adequate content * Some listener effort required

Table 2.5 (cont.)

Test of Spoken English Rating Scale for Overall Features

(ETS, 2002: 31)

30 Communication generally not effective: task generally performed poorly.	Speaker responds with much effort; provides limited speech sample and often run out of time. * Repair strategies excessive, very distracting, and ineffective * Much listener effort required * Difficult to tell if task is fully performed because of linguistic weaknesses, but function can be identified
20 No effective communication: no evidence of ability to perform task.	* Extreme speaker effort is evident; speaker may repeat prompt, give up on task, or be silent. * Attempts to perform task end in failure * Only isolated words or phases intelligible, even with much listener effort * function can not be identified

Apart from classifying the criteria according to the three groups of users, teachers can also score test takers' speaking skill using these following methods: objective scoring , holistic and analytic ratings (Bailey, 1998).

2.4.1 Objective Scoring

It is the way to score each test item with only one correct answer. Thus, it can be done by untrained scorers using a scoring key or by a computer via a computer-based test or scannable answer sheets.

2.4.2 Holistic Rating

It refers to the rating of an overall impression of learners' speaking performance with a single global rating (Mcnamara, 2000; Brown, 1996; Luoma, 2004). The advantages of holistic scales based on Luoma (2004) is that it is fast to rate from one score which includes both learners' strengths and weaknesses in the descriptors. However, it is not suitable for diagnosing the individual strengths and weaknesses. Focusing on less important points such as qualifiers or quality words, the descriptors do not identify the differences between ability levels explicitly. An

example of the test that uses a holistic scale is “The Finnish National Certificate Scale” (National Board of Education, 2002 cited in Luoma, 2004).Table 2.6 illustrates the version for the examinees and teaching professionals.

Table 2.6

The National Certificate Descriptive Scale

(National Board of Education, 2002 cited in Luoma, 2004: 61)

6	Speaks fluently with few if any non-native features, such as a foreign accent. Is capable of expressing even subtle nuances of meaning with precision, and also makes varied and appropriate use of idiomatic expressions. Is able to describe even a complicated topic and to include sub-themes in the description, to develop different viewpoints and to bring the presentation to an appropriate conclusion.
5	Speaks fluently without frequent obvious need to search for an expression. delivery characterized by naturalness, coherence and appropriate length. is able to present a clear and detailed description of even a complex topic. Can use idiomatic expressions and everyday expressions, and is able to express nuances fairly well
4	Copes fairly well even in less familiar speech situations. Makes a distinction between formal and informal registers, at least to some extent. Is able to present and justify an opinion comprehensibly. Is able to talk about or describe sights, sounds and experiences. Is obliged only rarely to use circumlocutions in everyday communication because of inadequate language proficiency.
3	Copes with the most familiar speech situations and is able to take the initiative in everyday language-use situations. Speech may be quite slow but there are few unnatural pauses. Is comprehensible despite transferring native or foreign language structure and vocabulary to the target language. Pronunciation may clearly deviate from target language standards.
2	Copes with routine speaking situations that require a simple exchange of information. Nevertheless, the speaker’s language proficiency considerably restricts the range of matters that can be dealt with. Successful communication of a message presupposes that the interlocutor is willing to help the speaker in forming the message. Pronunciation may deviate clearly from the target language norm, thus requiring special effort from the interlocutor and impeding successful communication.
1	Is able to ask or reply to simple questions dealing with immediate everyday needs. Can make use of simple polite forms. Copes with the very simplest speaking tasks, but communication is slow and very fragmented. Often obliged to resort to nonverbal means in order to be understood.

2.4.3 Analytic Rating

It is used to rate each of a number of aspects of students' language performance separately to show their language sub-skills. Analytical criteria include the performance level expected for each of the specified criteria based on the theory of what speaking is (McNamara, 2000; Brown, 1996; Weir, 2005). Due to the specified criteria, analytic scales provide detailed descriptors for raters to score students' speaking ability easily. For examinees, it can be used for the diagnostic purposes of their speaking performances. In addition, the information of a candidate's specific speaking performance will contribute in the course design (Luoma, 2004; Weir, 2005). In order to prevent a halo effect caused by each criterion to assess the same thing or overlap the others, Weir (2005) suggests that teachers write each level of language performance to be distinctively different. Likewise, the analytic method takes more time in scoring than the holistic one. Weir (1993 cited in Weir, 2005: 195) proposes an example of analytic speaking criteria for fluency as given below.

Fluency

4 Generally natural delivery, only occasional halting when searching for appropriate words/expressions.

3 The student hesitates and repeats himself at times but can generally maintain a flow of speech, although s/he may need an occasional prompt.

2 Speech is slow and hesitant. Maintains speech in a passive manner and needs regular prompts.

1 The students speaks so little that no "fluent" speech can be said to occur.

.

To choose between analytic and holistic scoring depends on not only the objectives of the test but also the circumstances of the scoring. Analytic scoring should be used if the diagnostic information is needed. With regard to the scoring circumstances, if scorers are less well-trained or they score in many places, an analytic scoring is suitable. But if the test is conducted by the raters who share the same experiences in attending the rater training or the test is scored in the same place, then "holistic scoring" is more essential (Hughes, 2003).

The other question raised before rating learners' speaking performance is how to analyze the data quantitatively or how to divide the transcribed data into units in order to assess language features such as accuracy, complexity or fluency. Though researchers define the units used in segmenting spoken data under the same name,

their details are categorized into three main groups, namely, semantic, intonational and syntactic. Among these three categories, the syntactic units gain more popularity than the others. Foster et al. (2000) divide syntactic units into 4 main units as follows:

1. Sentence : It is found that dividing the spoken data into sentences causes lots of problems.

2. Idea Unit (IU) : Kroll(1977: 90 cited in Foster et al., 2000) views IU as “a clause with pre and post v-clause, non-finite subordinate clauses as well as finite relative clauses” .

3. T-unit : T-unit following Hunt (1970 cited in Foster et al., 2000) and Young (1995 cited in Foster et al., 2000) refers to one main clause including any other subordinate clauses which relate to it. T-unit is the most used unit for analyzing spoken language data. According to Young (1995: 38 cited in Foster et al., 2000), the examples that are counted 1 T-unit are one main clause, one clause with subordinate clause, the 1-or 2-phrased apposition and any clauses in elliptical forms. These following elements are not t-unit : back channel cues like hmm or yeah and discourse boundary markers like okay, thanks or good.

4. C-unit or a Communication Unit : It occurred with an attempt to include elliptical answers into T-unit. Hence, C-unit in Lin & Hedgecock’s views (1988:45), includes elliptical answers in grammatical predication and are counted as a C-unit. For instance, the answer “Yes” of the question “Have you ever been sick?” is admitted as a C-unit (Loban, 1966 cited in Foster et al., 2000).

Implementing these units to analyze the spoken language data depends on what the raters put an emphasis on. For those which are clause-based, analyzing the sentence node and the idea unit will suit them. With regard to supra-clausal answers, analyzing C-unit and T-unit can be used. Compared with supra-clausal unit, clause-based unit can be analyzed more easily. But using supra-clausal unit offers greater validity in that the test-takers can produce more complicated speaking performance. Moreover, T-unit is accepted as “the most popular unit for the analysis of spoken language” (Foster et al., 2000: 360). Bygate (2001) defines T-unit as a finite clause and any subordinate clause dependent on it and implemented the unit to his study. In order to find the quality of the language features and accuracy, he calculated from the incidence of errors per T-unit, the higher the number, the less accurate the spoken language. For the complexity, he calculated from the number of words per T-unit, the higher the number, the more complex the language. The higher the number, the less

fluent the talk. Calculated from the number of unfilled pauses per t-unit, fluency was found out. From the sentence “When he arrived home, Alex took a bath and quietly ate the dinner alone.”, the accuracy score is 0 since there is no error in this T-unit and the complexity scores are 14 based on Bygate’s criteria (2001). Likewise, Teng (2007) also used T-unit in her data analysis with different criteria. To analyze the quality of accuracy, the number of error-free clauses are calculated as a percentage of the total number of clauses in a given speech samples. For complexity, she measured it by dividing the number of clauses by the number of C-unit. To calculate the fluency value, she divided the number of syllables by the seconds to produce the spoken language.

In this study, the analytical rating scales of accuracy, fluency, complexity, and comprehension were selected in order to get the results to improve students’ oral abilities or to be used for diagnostic information. The criteria were written in the form of descriptors instead of measuring by using T-unit since T-unit only gives the results quantitatively or in numbers which cannot describe students’ strengths and weaknesses. Furthermore, T-unit cannot give a full analysis of spoken discourse (Foster et al., 2000) and is not suitable for the utterances with a few complete sentences with so much hesitation and repetition (Tarone,1985).

2.5 Task-based Language Learning

2.5.1 What is a task?

For the purpose of increasing the number of good communicators in English, the governments especially in East Asia have changed the educational policies to increase the numerous versions of Communicative Language Teaching (CLT) (Littlewood, 2007). The countries in Asia-Pacific, for example, implement CLT particularly Task-Based Language Teaching (TBLT) which is “the latest methodological realization of CLT” (Nunan, 2003: 603) in their educational policies and pedagogic principles. Though countries in this region including Thailand and the Philippines may not mention TBLT formally in their syllabuses, TBLT is integrated in the implementation of their ELT curricula and learning materials (Todd, 2006; Vilches, 2003 cited in Littlewood, 2007).

In a TBLT classroom, the teacher acts as a facilitator while the student as a doer and what the student has to do is to discover the spontaneous and authentic language uses (Oliveira, 2004). In order to understand the concept of TBLT clearly,

what the task really means should be first considered. The definition of the word “tasks” vary depending on the ways in which they are used. Williams and Burden (1997) claim that a task is the activity in which language learners participate in order to pursue their language learning. Bachman and Palmer (1996) have defined “a language use task” as an activity which provides the opportunities in using language in order to attain a specific goal or objective in a specific situation. Similarly, Willis (1996) views tasks as a series of activities to do towards the outcome by using the target language. While a task following Bygate, Skehan, and Swain (2001) is an activity which the test takers try to achieve the goal by using the target language and the focus is on the meaning, Norris, Brown, Hudson, and Yoshioka (1998 cited in Bachman, 2002) extend the scope of the activities to those done in daily life achieved by using the target language. Tasks for all of the theorists mentioned earlier are more concerned with the language activities using the target language to achieve the purposes by mainly focusing on the meaning. Likewise, Estaire and Zanon (1994) emphasize meaning rather than form to distinguish the “communication tasks” from “enabling tasks”, in which the main focus is on linguistic aspects (grammar, vocabulary, pronunciation, functions, and discourse). Their ‘enabling tasks’ would correspond to Ellis’s (2000) term “exercises” in which the learners have no communicative purpose and be in contrast to “communication tasks” .

Some try to distinguish between two kinds of tasks: real-world or target tasks (uses of language in the world beyond the classroom) and pedagogical tasks (use of language occurring in the classroom) (Nunan, 2004). Long (1985: 89 cited in Ellis, 2003) presents a target task in a broader view that “ a task is any pieces of work done in daily life, at work, at play and in between”. The task in Long’s view relates to what they will do in the real situation though it produces non-technical and non-linguistic outcomes or sometimes does not include language use (Nunan, 2004). In contrast, pedagogical tasks according to Ellis (2003) refers to the language process that the learners try to achieve the goals by using their linguistic abilities in both producing and receiving the target language. Nunan (2004) refers to pedagogical tasks as the way learners use the target language to comprehend, manipulate, produce or interact in classroom work with the emphasis on meaning rather than on form. Likewise, Prabhu (1987) claims that it is an activity which the learners go through the cognitive process to achieve the outcome. In short, these definitions embrace “the process of thought” which is the crucial factor in performing tasks. Besides, tasks include communicative language use and the meaning focus. This main characteristic is in

accordance with one of the characteristics of the task cited by Skehan (1998: 95) as follows:

A task is an activity in which

- meaning is primary;
- there is a communication problem to solve;
- there is some sort of relationship to comparable real-world activities;
- task completion has some priority;
- the assessment of the task is in terms of outcome.

Many educators mentioned earlier define tasks in general perspectives. In order to respond to the different purposes of using tasks, the definitions of tasks need to put weight on each language function differently. Thus, Bygate et al. (2001: 11-12) propose the idea of "all-purpose definition" which can be adapted to other language areas such as pedagogy, pragmatics or testing. The following definition is the most basic one from them.

> *A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective.*

Then they give the example of how to adapt the above definition to make it more direct to other areas like the research perspective on testing and assessment as follows :

> *A task is a contextualized, standardized activity which requires learners to use language, with emphasis on meaning, and with a connection to the real world, to attain an objective, and which will elicit data which can be used for purposes of measurement.*

2.5.2 Task – based Language Assessment

Tasks are also implemented in language assessment as "Task-based Language Assessment" (TBA) and diverse definitions are emerged to fit their uses. Oliveira (2004) refers to TBA as the method the students from a task-based instruction class express what they have learnt. These following meanings focus on real world situations that give the test takers a chance to use authentic languages. Elder et al. (2002), for example, term task-based assessment as the assessment of learners when they perform real-life tasks with simulation of the situations that let them use the "authentic sample of language". Likewise, Ellis (2003) views task-based assessment as a tool for eliciting and evaluating communicative performance of the learners in the context of language use, that is, meaning focused and directed toward some specific roles. Norris et al. (1998) also include authentic tasks as crucial characteristics of performance assessment, a broader approach to assessment including task-based testing. Besides, they add two more characteristics of the

assessment, namely the performance assessment must be based on tasks and only qualified raters will decide whether the task is successful or not. Apart from the ability in use as in the above-mentioned definition, Brinley (1994 cited in Bachman, 2002) includes language knowledge in assessing the learners' language proficiency. In addition, task-based language assessment is cited to embrace all the testing process which the main focus is on the test takers' language performance as follows :

> *Task-based language assessment takes the task itself as the fundamental unit of analysis motivating item selection, test instrument construction, and the rating of task performance. Task-based assessment does not simply utilize the real world task as a means for eliciting particular components of the language system, which are then measured or evaluated; instead, the construct of interest is performance of the task itself."*
>
> Long and Norris (2000: 60)

Though the definitions of tasks do not directly mention what language skills are used in doing tasks, numerous researches on tasks assumed that tasks are directed at oral skills, especially speaking which may involve other three language skills, i.e. listening, reading, and writing (Ellis, 2003). Based on "Language Use Task" of Bachman and Palmer (1996), "speaking tasks" in Luoma's view (2004) are the activities that have speakers participate in goal-oriented language uses. Thus, speaking test tasks should include the activities that allow the test-takers to use the target language in achieving their goals in communicating in different situations.

2.5.3 Tasks for second language oral tests

Due to the realization of the effectiveness of tasks toward language pedagogy, tasks are included as the basis of syllabus design as well as classroom activities (Skehan, 2001). Since tasks can elicit learners' language use in actual language situations, tasks are not only salient to language learning but also to language testing as Wiggleworth (2001:189) mentioned that

> *"While task-based curricula lay claim to providing more authentic interactions within language pedagogy, the use of tasks in test design is a reflection of a move towards a more communicatively oriented approach to language assessment."*

Though tasks are useful in many educators' views, there are some criticisms about adapting them in an achievement test. A goal-oriented not rule-oriented method, tasks may

not be assessed systematically as the language structure designed in a curriculum. As a proficiency test, test tasks may have some problems in generalizing from the language competence to the real use of language ability in different contexts and conditions. To compensate for these weaknesses, many researchers have tried to find how to design the speaking test based on the models that can include more language elements and can be the frameworks from which the oral performances can be inferred.

Skehan (1998) is one of the investigators who proposed numerous factors influencing test takers' oral performance shown in Figure 2.2 (adapted from the model of Kenyon and Mcnamara (1992; 1996 cited in Skehan, 1998). His model includes two more variables, namely task components and the candidates' ability for use.

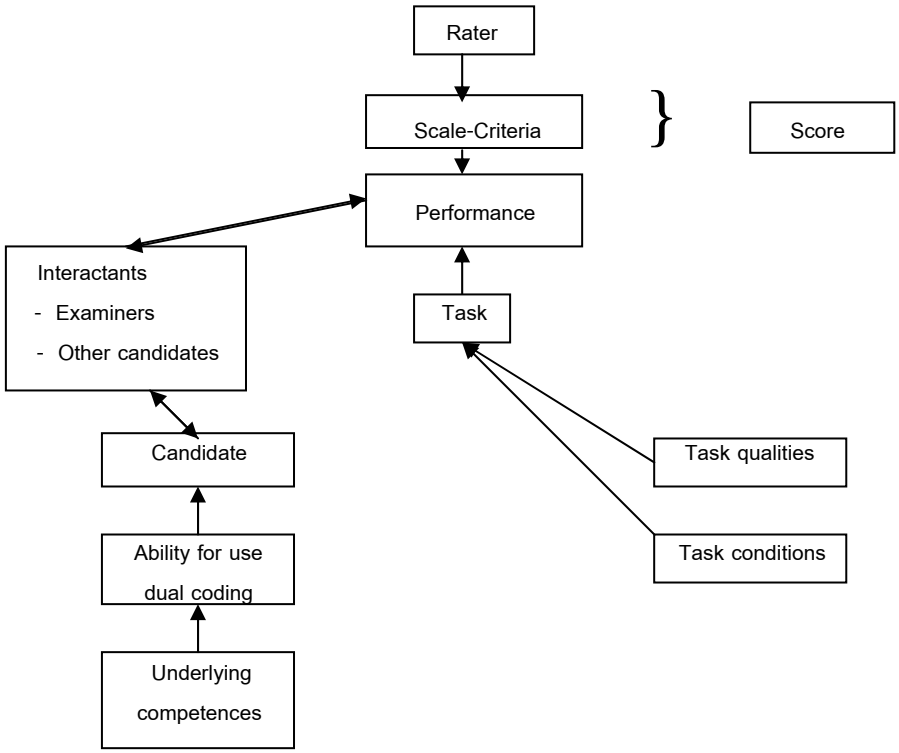


Figure 2.2 : The model of oral test performance by Skehan (1998: 172)

The performance influenced by numerous factors is placed in the centre of the model. The task is used as the vehicle for the performance and raters are responsible for judging the performances based on the scale criteria in terms of the scores. The performance is also affected by the interactants that involve the examiners (when they participate in a traditional oral interview) as well as other candidates (when they attend "group-based oral assessment"). The ability of the candidate is the major goal of the actual assessment. Thus, assessing learners' oral abilities requires not only an assessment of

"underlying competence" but also "the ability for use". The former is one of the components of learners' communicative competences. It affects the task, language performance and particularly the assigned score. The latter is a set of abilities relating to the underlying competences with actual performance in real situations. Since having underlying competences is not so important as capacity to transform these competences to arrange language system into the proper order so the learners perform the languages well. Candidates' ability for use; therefore, will act as the strategic competence in making the underlying competences meaningful. Then the candidate will choose the performance in responding to the tasks. With regard to tasks, they have impacts on the prediction of performance and the generalization across context. Tasks are influenced by two factors: task conditions (the conditions in which the tasks are administered) which relate to how learners' performance can be predicted and task qualities (task characteristics or the features of the tasks) which connect to the generalization across contexts (Skehan, 1998; Wiggleworth, 2001; Skehan, 2001).

Another factor to be taken into account when designing the test is two kinds of language tests or whether the test is direct or indirect. A direct test is the test that assesses criterion performance directly or the test performance is similar to the criterion performance. Test takers' performance is assessed holistically depending on "a direct sampling of the target performance" (Ellis, 2003). A direct test is ; therefore, the real-world performance that learners need in communication (Nunan, 2004). Robinson and Ross (1996: 460) view an "indirect test" as "the incomplete versions of the target criterion procedure". Only the particular language components are focused on in each test. Then the test taker's performance is assessed analytically (Ellis, 2003).

Moreover, two methods of language tests that also achieve the task-based assessment according to Baker (1989) are system-referenced tests and performance-referenced tests. System-referenced tests do not aim to assess language proficiency in specific language use but only in general including a phonological, lexical or grammatical system of the language. The purposes of the tests are to evaluate language mastery as a psychological construct without specific reference to any particular language use. Thus, they can be inferred to different test samples and are easy to design and practical. But their weak point is that they lack face validity because they cannot respond to "the holistic procedural skill" which is the aim of most language teaching programs. Moreover, the test can assess only the achievement of procedural and communicative purposes which test

only the discrete knowledge of the language system; therefore, the test also lacks construct validity (Robinson and Ross, 1996).

In contrast, as a content-oriented test, the performance-referenced test intends to assess the ability to use the language in specific contexts. Consequently, it has greater face and construct validity in the way in which the future language abilities of learners can be predicted from the test scores obtained from the tests which simulate the real-world activities. However, this type of test also has disadvantages in the generalizability problem since the test is adapted for some test takers' specific needs. In addition, the simulation activities make this test not practical due to the complicated details of simulations which has to be closer to the future language performances (Robinson and Ross, 1996). The two main methods of language tests tasks result in the intersecting dimensions as shown in the following table.

Table 2.7

Test Referencing and Test Directness

(Ellis 2003: 285)

more general ← → more specific

more analysis ↓

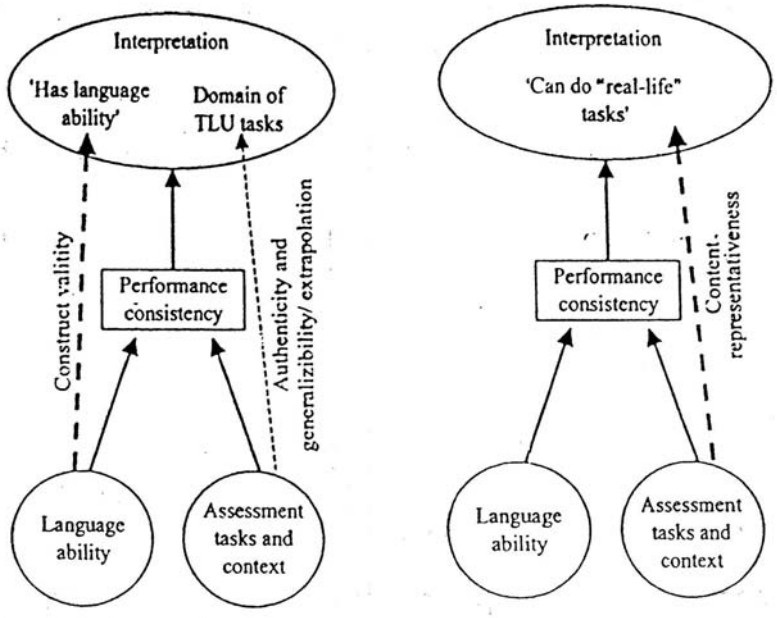
Mode	System-referenced	Performance-referenced
Direct (holistic)	Traditional tests of general language ability : Samples of oral or written language - oral interview - free composition Information transfer tests: - information-gap - opinion-gap - reasoning-gap	Specific-purpose tests : - tests based on observing real-world tasks - communicative simulation of target tasks e.g. Library Skills Reading Test(LSRT)
Indirect (analytic)	Discrete-item tests of linguistic knowledge : - grammar, vocabulary and reading multiple-choice tests - elicited imitation of specific linguistic features - error-identification tests Integrative tests : - cloze - dictation	Tests that seek to measure specific aspects of communicative proficiency discretely : - tests of specific academic sub-skills e.g. the ability to cite from a published work - tests of the ability to perform specific functions or strategies, e.g. the ability to write a definition of a technical term.

For direct system-referenced tests, the language samples are elicited in order to allow language learners perform language skills which can be examined and analyzed depending on the “language component part”. This kind of test includes traditional tests, that is, the tests which are based on communicative tasks like transferring information. Though they are difficult to assess the performance, they are practical in administration (Baker, 1989).On the other hand, indirect system-referenced tests aim to assess specific knowledge instead of the criterion performance. They are integrative language tests with

multi-test items. The direct performance-referenced tests focus on the language abilities for specific purposes and they simulate what test takers have to perform in the real world tasks. The test performance is related to the criterion performance. For instance, if the criterion performance is using an English-English dictionary to solve problem in reading, the students who can perform well in the "how to use a dictionary" test by identifying the correct meaning of the words in the sentences will also use a dictionary in the future (Baker, 1989). Designed analytically, indirect performance-referenced tests involve performances of specific skills or abilities to perform specific functions. The test performances are not similar to those of criterion, namely, the performances taken from the criterion have to be analyzed and to be in the abstract form.

We can see that the indirect assessment including discrete point tests is more convenient in the way that it is administered and rated. On the other hand, the direct assessment especially the performance-referenced test like speaking tests is difficult to design and to be scored. Despite its inconvenience to testers, a direct performance–referenced test is recommended to be used to predict learners' abilities in real-life language uses in the future (Norris et al. 1998). As one kind of performance assessment, task-based performance assessment involves not only real-world tasks used in the testing process but also language ability assessment and performance emphasis.

It is known that the primary purpose of language testing is the inferences that one can make from the test scores to learners' language ability (Bachman and Palmer,1996). Moreover, the scores reflect the contents of the test tasks which are the representatives of the target language use domains outside the test. Brown et al. (in press cited in Bachman, 2002) are interested in inferring language performances from the test scores. Figure 2.3 illustrates two different ways of test interpretations: (a) ability–based inferences and (b) task–based inferences. The way to design ability–based tests or construct–based tests is that the test should be on the basis of performance quality of the test construct, score uses and scoring criteria. For the ability-based inferences, the test scores are the indicators of learners' language abilities according to the language constructs the testers design the test on and the scores are also generalized to the target language use domain. Developing task–centered tests depends on "the desired performances" and part of the performance tests include score uses and scoring criteria. The scores from this kind of test can be inferred to predict the learners' future language performances on real world tasks or what the test takers are able to do (Norris et al., 1998).



a. "Ability-based" inferences about language ability ***b. "Task-based" predictions about future performance as "real-world" task***

Figure 2.3 : Different interpretations of response consistencies on language assessment task (Bachman, 2002: 457)

In summary, the construct–based tests emphasize constructs and tasks but the task – based tests emphasize only performances on tasks (Bachman, 2002). The performance-referenced test concerns with "what is done with language" and "it owes their development to the desire to have information about what a testee can actually do with his language proficiency (Baker, 1989: 7) while the system-referenced test assesses the candidate's language ability to control certain language structures and the size of his vocabulary or emphasizes "a code to be mastered" (Baker, 1989: 10). Unlike the system-referenced test which can assess only the language construct of the course, the CMLSTT, a performance-referenced test, is aimed to assess the learners' abilities in achieving language functions. Therefore, the obtained scores will not indicate the language ability the candidate possesses but the language functions the learners can successfully perform in their future real life situations

2.5.4 Language Test Task Characteristics

Another factor that should be taken into account apart from different types of test tasks is characteristics of the test tasks. Applied linguists' opinions toward the characteristics of the test tasks vary in components they pay more focus on. In order to

design the speaking tasks apart from the models of language ability mentioned previously, task characteristics are another factor that should be taken into account. Candlin (1987 cited in Nunan , 2004: 40) mentions the following task characteristics :

Input : *the data presents for the learners to work on*
Roles : *the relationship between participants in a task*
Setting : *where the task takes place: in class or out of class arrangement*
Action : *the procedures and sub-tasks to be performed by the learners*
Monitoring : *the supervision of the task in progress*
Outcomes : *the goals of the task*
Feedback : *the evaluation of the task*

Nunan (2004) adds some more characteristics by proposing the following diagram of the task components which include at least three elements: goals, input and procedures. These elements are supported by roles and setting as shown in the following figure.

Figure 2.4 : The characteristics of the task (Nunan , 2004: 41)

Goals refer to the reasons supporting why the students have to perform the task. They also relate the task to the learning curriculum. Goals may not mention directly how teachers and learners behave but they can be inferred from the tasks. Goals may not relate to language but other subject areas such as communicative, sociocultural or process-oriented types. The most important concept is that goal statements should be related to the students in terms of their observable performances. With respect to input, it includes “the spoken, written and visual data that learners work with in the course of completing a task.”(Nunan , 2004: 47). Good input should be simplified to help learners process the language more easily. Increasing the frequency of the target language makes the learners familiar with the language forms and rules. Moreover, slowing down the speed of the speech makes it easier. Using authentic texts which are suitable for the learners’ needs, interests and proficiency levels helps their understanding. Apart from the authentic input, the procedures which specify what the learners would do with the task should be authentic.

The procedural authenticity should simulate what the learners are expected to use the language skills in communicative interactions outside of the classroom. Regarding role, it refers to “what part the learners and teachers are expected to play in carrying out learning tasks as well as the social and interpersonal relationships between the participants.” (Nunan, 2004: 64). Learners can be both active and passive recipients depending on the language activities. In order to stimulate the learners to have more active roles, teachers have to adopt different roles as facilitators, participants as well as observers and learners. Any parts the teachers play depend on the roles that are appropriate and are needed by the students. The last element, setting, it is the arrangement of the classroom specified in the task. There are two things to take into account when considering settings: mode and environment. Whether the students conduct the task individually or in a group refers to “learning mode”. “Environment refers to the place where learning takes place” (Nunan, 2004: 72). These places include the classroom, the language center as well as technological learning environments such as web-based instruction, the internet or cable television.

Chalhoub-Deville (2001 cited in Ellis, 2003: 287) proposes three main characteristics of oral tasks as follows :

1. They must reflect learner-centered properties, that is, tasks have to support learners to express themselves by using their background knowledge and experiences.
2. They must be contextualized tasks which stimulate learners to use cohesion and coherent discourse and be similar to real-world situations.
3. They should be authentic in the real-life use, i.e. the language used in the test is closely related to that in the real world.

According to Bachman and Palmer (1996), the language test task characteristics are the same as those of Chalhoub-Deville (2001) in that they emphasize learners’ participation and add two more features: a particular situation and a goal-orientation. Among the characteristics of the test tasks mentioned above, the most popular one is that of Bachman and Palmer (1996) which includes five aspects of tasks as follows :

1. Characteristics of the setting refer to the testing environment, the test takers’ familiarity of materials and equipment, participants as well as administration time.

2. Characteristics of the test rubrics include the test structure, instructions, time allotment and scoring method.

3. Characteristics of the input or materials put in the test task are both format and language characteristics.

4. Characteristics of the expected response refer to two kinds of responses: selected and constructed response which can be limited or extended answers depending on the length of the response.

5. Relationship between input and response is in three different aspects: reactivity, scope and directness of the relationship.

2.5.5 Studies on the effects of task characteristics and task conditions

Foster and Skehan (1996) studied whether three task types (personal, narrative and decision) had a different impact on three language components: accuracy, fluency and complexity. The personal task was concerned with the information that was well-known to the participants, i.e. the personal information exchange task. In the narrative task, the students had to construct the story from a series of pictures with no obvious storyline. Students were required to take the role of judge in the decision-making task. They found only a few effects on students' language performance. Regarding the personal task, learners generated more fluent languages but less complex than the other two tasks. The language students used in personal and decision tasks had higher accuracy than narrative tasks. Not only the characteristics of task types influence students' language performance, it was also found that the interaction between planning time and the characteristics of task types affect complexity and fluency. In narrative and decision tasks, the abilities in complexity and fluency improved but not much in the personal task.

As a series of the studies, Skehan and Foster (1997) researched the same kinds of task types but with different situations from the previous work. Students produced the least accuracy in narrative tasks. The language complexity in personal and decision tasks increased while the narrative task had the least complexity. The narrative task generated the most fluent performances while the least one was the personal task. The result of the study in 1996 was different from that in 1997 since the language performance was also influenced by the nature of the tasks. They concluded that the more-structure form of the personal task (1996), namely 'telling the way back home' generated less language complexity than the free-form one such as the personal task (1997), i.e. 'expressing our own ideas'. While the well-structured tasks help create the language fluency, doing tasks that have to retrieve information or propose the solutions needs more time so they generated less language fluency.

Skehan and Foster (1999) investigated the influences of task structure and processing load on performance on a narrative retelling task. 47 low-intermediate students

did one of the two tasks (structured narrative called restaurant task and unstructured narrative named golf task) under one of the four conditions (from the most cognitive demanding to the least) as follows:

1. watch and tell simultaneously
2. storyline given, watch and tell simultaneously
3. watch first, then watch and tell simultaneously
4. watch first and then tell

In terms of the two task types, the structured task (restaurant task) produced more fluent language performances than the unstructured one (golf task). The differences in performances across the two tasks in complexity and accuracy were hardly found. With respect to four task conditions, the fluency scores in “watch and tell” and “watch then tell” conditions have the most similarity of the high (disfluent) scores. As for the highest complexity and accuracy scores, it was in the “watch then tell” condition. The accuracy scores under “the storyline, watch and tell” condition was also at the high level. Using the two-way ANOVA, there are a significant task effect for fluency, a significant condition effect for complexity and a significant interaction effect for accuracy. The restaurant task generated more fluent scores as the task including clearer sequential structures resulted in more language fluency. This is because the candidates could focus directly on the speech via the accessible macrostructure of the task instead of wasting time doing the mistake repair of the golf task. The complexity scores under “The watch, then tell” condition were the highest due to a non-simultaneous condition while the simultaneous conditions caused less complexity. Regarding accuracy, there were no significant main effects either from task types or test conditions but the combination between them showed a significant result. The major impact was the interaction of the structured task and the “watch then tell” condition which produced the highest accuracy. The planned condition as pre-task preparation in combination with the structured task had more impact on language accuracy while the unstructured tasks under this condition generated a low accuracy value. The structured task under most task conditions had more accurate language performance than the unstructured one except for the unstructured task in “the storyline, watch and tell condition”**.**

Skehan (2001) studied the effects of five test characteristics toward examinees’ language performances as shown in Table 2.8.

Table 2.8

Summary of the Effect of Task Characteristics on Complexity, Accuracy and Fluency (Skehan 2001: 181)

Task Characteristic	Accuracy	Complexity	Fluency
1. Familiarity of information	No effect	No effect	Slightly greater
2. Dialogic vs. monologic tasks	Greater	Slightly greater	Lower
3. Degree of structure	No effect	No effect	Greater
4. Complexity of outcome	No effect	Greater	No effect
5. Transformations	No effect	Planned condition generates greater complexity.	No effect

Since tasks compose of the information that is part of their personal experiences, "the familiarity of the information" refers to the retrieval of relevant information as the basis for completing the task. This study includes both dialogic tasks such as decision making tasks and monologic tasks such as narrative and personal tasks. Degree of structure and complex outcomes also vary from one task to another according to the nature of the tasks. Some tasks can be completed with a single idea while the personal judgement together with other learners' ideas are needed in others such as in decision making tasks. Transformation of task material means to retrieve the information or judgements about presented materials. Since this study was extrapolated from the studies carried out in the classrooms and laboratories, the researcher did not plan to compare the three language features (accuracy, fluency and complexity) of five test characteristics at the beginning of the study. The number and the nature of the tasks are not systematic; therefore, the effects of a number of task characteristics cannot be found. As a result, the generalization of this study can be only tentative. In addition, the researcher cannot conclude that the differences of students' language performances from doing diverse tasks result from their real language abilities only or also from the task characteristics. Consequently, he suggests systematically organized studies and other further studies that use the same features of the test. Moreover, researchers should be careful with inferring from one language

performance to another. Finally, the test task should be arranged depending on the aspects of performance; that is, accuracy, complexity and fluency according to their importance.

Similarly, Wiggleworth (2001) conducted a study including four factors, that is, two test characteristics (structure/non-structure and task familiarity) as well as two task conditions : native speakers/non-native speakers(NS /NNS) and planning time which might influence learners' language performances. Similar to the results of Skehan's (2001), there were no significant differences in learners' performances. Although the differences across the tasks were identified, no interaction effect was present. She concludes that different factors influence different types of tasks to different degrees but it might not be strong enough to be assessed. Teachers have to design the test strictly to the parameters and trial with a range of learners and interlocutors.

2.5.6 Studies on the effects of task difficulty

With the belief that the harder tasks will yield less language fluency but more complex and accurate scores than the easy ones, Robinson et al.(1996 cited in Skehan, 1998) compared the differences between the hard task (a "there and then" narrative) and the easy one (a "here and now" task). They found that the hard task without any immediate visual support, that is, cartoon strips increased a processing burden for students. They resulted in more complex language due to the fact that students had to concentrate only on the speech. But it produced lower fluency.

Elder et al. (2001) investigated three design variables in picture narration tasks in semi-direct speaking tasks, i.e. immediacy (here-and-now vs. there-and-then tasks); adequacy (telling vs. inventing a story); perspective (narrating in the first or third person). The test takers' performances were rated by external evaluators and the discourse analysis was done in order to find out whether the manipulation of task instructions, materials and performance conditions would have any systematic effects on the fluency, accuracy and complexity of the test takers' speech or on examinees' perception of task difficulty. However, no significant effect was found in their language performances or their perception of difficulty with the exception of the here-and-now task, which they produced more accurate language. This result contradicts the finding of the previous study conducted by Robinson et al. (1996).

2.5.7 Studies relating to the effects of task types

In addition to the task characteristics, other researchers especially from the EFL countries put the emphasis on the effects of task types. Teng (2007) investigated the effects of three different task types: answering questions, picture description and presentation on Taiwanese college students' speaking performance. Though the students' scores were not significantly different on holistic measures, the complexity of the answering question task was better than that of the picture description. Moreover, the students spoke more fluently in answering the question task than in the other two tasks. According to Skehan & Foster (1999 cited in Teng, 2007: 8-9), "Complexity of language was influenced by processing load and the conditions under which tasks were done, especially related to the processing demands that they entailed. There seemed to be more processing load for answering questions in this study since test takers were expected to provide answers directly related to the questions. As for picture description, less processing load was involved because of the flexible nature of description. In terms of fluency, answering question task seems to be the most common EFL speaking tests and activities for Taiwanese college students. It is possible that the subjects can speak more fluently in answer question tasks than in the other two tasks". In addition, subjects perceived the answering question task as the most stressful while describing pictures was their favorite and suggested it be included in the speaking tests.

Weir and Wu (2006) studied three forms of the General English Proficiency Test Intermediate Speaking Test (GEPTS-I) containing three different task types : read-aloud, answering questions and picture description. The three tasks are statistically parallel and Form two and Form three are parallel at the individual task level. The findings of this study shed light on further improvement in parallel form reliability.

2.5.8 How to assess performance in task-based assessment (TBA)

Following Ellis (2003), task-based assessment (TBA) not only provides a measure of the examinees' language ability but also elicits their performance. In order to assess the performance in a task-based test, there are three principal methods available: direct assessment of task outcomes, discourse analytic measures and external rating. The details of each method will be described and discussed below.

1. Direct assessment of task outcomes

The outcome of the test can only be assessed with a closed task that results in a solution that is either right or wrong. Assessing tasks directly helps to afford an objective

assessment without the judgment on the part of the assessor involved. Also, this test is easy and quick to measure. However, this method might be difficult to administer since the tester has to observe each candidate as he/she completes the task. This difficulty could be overcome if the task requires a written product which can be assessed after the task has been completed. Moreover, it is not clear to what extent direct assessment measures language ability as opposed to non-linguistic abilities or general knowledge, especially direct performance-referenced tasks derived from a work-sampling approach.

Based on the concept of direct assessment, an oral test technique called “Making Appropriate Responses” is proposed by Underhill (2003).The test involves situations in written or spoken forms which are designed to be easily understood. The test-takers have to assume themselves to be in each situation and respond to one or two appropriate spoken sentences. The test is designed to elicit functional language which is accepted as a crucial part of language proficiency. The marks on a simple scale are used for this test such as

0 for inappropriateness or seriously incorrect sentences

1 for relevant but not entirely acceptable sentences

2 for appropriate and correct sentences

The answer of the test-takers will be compared with the key answers prepared from trying each situation out, but it is rather difficult and impossible to include every acceptable answer. Despite the convenience and short period of time in administration and marking, there are still some points which should be taken into consideration :

- The situation should not be ambiguous and there should be only one answer.
- The cultural background of the test-takers should be the same as those from the trying out.
- Language variations may affect some students.

The purpose of marking keys of this technique is to save time and uncertainty by specifying the marking guideline to tell how markers should assign the marking of each question or task. A comprehensive marking key aims to suggest the way to deal with the problem the markers might face and identify the most important language areas as a guideline in marking unusual responses. Thus, the effectiveness of the marking key depends on the answers in the pilot study and the answer should be revised and discussed.

2. Discourse analytic measures

This method provides counts of specific linguistic features occurring in the discourse resulting from performing the task. The assessment varies and it could focus on the candidate’s linguistic competence, sociolinguistic competence, discourse competence,

and strategic competence. The purpose of using this method is to afford a more or less objective measure of the specific aspects of the examinees' performance that have been chosen for assessment. This method; however, is not generally used in TBA. It is just used as a means of comparing the assessments afforded by discourse analytic methods and external rating.

3. External rating

External rating involves the assessor observing a performance of a task and making judgment like in the direct assessment of tasks. The judgment, however, is more subjective. Rating scales are the most common method of assessing performance on task-based tests of both system-referenced and performance- referenced kinds. These scales specify the competency and the level of performance. A checklist of competencies, however, is used as an alternative.

2.6 Computer-assisted Language Assessment

Learning language through technology has played an important part especially in second language acquisition since the 21st century. Computer-assisted language learning (CALL), for example, is used in learning and teaching oral skills, listening and speaking in that it allows students to use language cooperatively in authentic situations (AbuSeileek, 2007). Due to the advent of communication with the computer or with other people via the computer, communicative competence has changed to refer to meaning transfer between the information technology and real world. The concept of communicative competence extends to the changes in our daily lives caused by this technological innovation. Thus, communicative competence for the learners is the communication related to electronic communication. Since the advances in information technology have been developed quickly, the computer hardware and software tools have grown and their costs have become cheaper. The use of computer for language testing is increasing. Language teachers in this century should adjust the information technology in language assessment in order to provide "the computer-using experience" into learners' language learning experiences and encourage the students to be ready for computer-assisted tests that they certainly experience such as IELT, TOEFL and DIALANG which is used for diagnostic purposes (Alderson, 2000 ; Brown, 1997).

Brown (1997: 45) views computer-assisted language testing as "the tests that are administered at computer terminals or personal computers." Computers ; therefore, play crucial roles in almost every testing process ,namely test construction, test delivery,

response analysis and score reporting (Chapelle, 2001). Furthermore, not only receptive-response test items such as multiple choices, true-false and matching items can be adapted to computer-assisted testing, interactive testing such as role plays, interviews or presentations can be possible in computer-assisted testing (Brown, 1997). In addition, computer-based language testing can compensate the limitations of paper-based language testing in that the multimedia, graphic and oral information can be presented to the examinees in a very realistic way. Simulating real-life settings can strengthen tests' authenticity and validity. The study of Choi, Kim and Boo, (2003) support the idea that computer-based language testing (CBLT) can be highly comparable to paper-based language testing (PBLT) in terms of content and linguistic features. Moreover, the constructs measured by the two tests are the same. Apart from the comparability between CBLT and PBLT, the test takers have positive attitudes toward CBLT and their attitudes toward CBLT compared with other types of oral proficiency including PBLT are not significantly different. (Brown,1997; Kenyon & Malabonga, 2001; Stricker, Wider, & Rock, 2004).

The advantages of using computer in language testing can be further subdivided into two categories: testing considerations and human considerations (Brown,1992 cited in Brown, 1997; Alderson 2000: 595)

1. Testing considerations cover the following advantages :

1.1 Computers are more accurate at scoring selected response tests and also at reporting scores.

1.2 Computers can give descriptive feedback immediately in the form of scores and basic testing statistics.

1.3 CBT supports the use of tests for low-stakes purposes like diagnostic tests which can give feedback on incorrect answers.

1.4 Tests delivered by the internet will give more updated information than in CD-ROM or diskette-based tests and test results can be sent immediately to the score users.

2. Human considerations include the following advantages :

2.1 The students can be individually administered. There are no needs for dealing with group –administered test and all of the organizational constraints.

2.2 CBT can be taken at any time that is convenient to the test takers.

2.3 There are no human proctors to wait for the students to finish the test.

2.4 The students do not feel frustrated with the overwhelming test items since the questions in the computer are presented one at a time.

However, the computer-based testing has disadvantages which can be grouped into two categories: physical considerations and performance considerations.

1. In terms of physical considerations, the following are some advantages :

1.1 The software or other tools are not available and are difficult to be implemented for some item types.

1.2 Screen size, font of the letters, the materials and graphic capacities are limited for some kinds of tests.

2. Some disadvantages of using CBT in performance considerations are as follows :

2.1 The differences of students who are and are not familiar with using computers can be troubles.

2.2 The test takers' anxiety can affect test performances.

2.3 Some types of language tests may be better if they are presented via paper-and pencil format.

With regard to language assessment, computers are used mostly in receptive skills. The study of Madsen (1991), for example, was a computer-adaptive reading test as a placement test for an intensive ESL program. Like other adaptive language tests, the number of questions and time allotment vary according to the different test takers. Apart form the studies that put more emphasis on designing computer-based tests, some researchers are interested in implementing the computer in their language learning programs and also use the computer-based test as an achievement test of their programs, such as the speech test in the format of the Discourse Completion Test (DCT) of Chiu, Liou, & Yeh (2007***)***. This test is used as a pre-test and post-test to measure students' learning speech act through a web-based conversation environment called "Candle Talk". The students have to respond by recording their answers through the microphone in the computer. The Non-English major group improves significantly more than the English major group from learning through the Automatic Speech recognition (ASR) -supported instruction. Moreover, most of the students agree that the instruction with ASR and the format of simulated real-life conversation could facilitate their speaking.

Conclusion

Assessing speaking has been changed from inferring the students' speaking abilities from the scores of discrete-point tests to students' participating in real-life situations. With a large number of students but a few administrators, assessing students' speaking performance directly such as an interview or conversation seems to be impractical. Semi-direct speaking testing compensates for the limitations of direct testing by having students interact with the computer instead of interlocutors. Though semi-direct speaking test is argued that its validity is less than that of direct speaking test, it can be used in some situations direct testing is impractical (Shohamy, 1994). Moreover, semi-direct speaking test has its advantages in administration and in comparison with other tests of the same constructs.

In order to design an effective oral abilities test, theoretical frameworks should be considered. The model of language ability proposed by Bachman and Palmer (1996) is chosen for Computer-mediated Semi-direct Listening-Speaking Test Tasks (CMLSTT). Only the knowledge of pronunciation, syntax and phonology are selected to be studied and used in the band descriptors for accuracy, fluency, complexity and comprehension. Regarding the studies on TBLT, task characteristics can affect students' speaking abilities. Thus, in designing the test, task characteristics should be taken into consideration. The task type is another factor that EFL studies focus on since it includes language activities and language functions as the test contexts which are the main purposes in conversing with people and are used as the framework for many formal speaking tests including TSE. In this study; therefore, three language activities: answering questions, describing pictures and transferring information as well as one language function (giving personal information) are chosen based on the course description of "Communicative English I Course" at Buriram Rajabhat University (BRU).

Since this study aims to design a speaking test to serve a number of first–year English-majored students at BRU, a direct speaking test is not suitable for this purpose. Computer-mediated Semi-direct Listening-Speaking Test Tasks might be an alternative to deal with the problem of a few interviewers. In addition, it is a good opportunity for the students to try a new testing approach which simulates real life situations.

Chapter III

METHODOLOGY

This chapter presents the procedures and the methodology of the research study which intended to determine the effects of three different task types and three groups of students on students' oral abilities and also to study the test takers' differences in four aspects of language performance in each task type.

3.1 Research Approach

This study is descriptive research focusing on the CMLSTT test development. The study aims to compare the oral abilities in terms of three different task types as well as the three groups of students and investigate whether these three task types and the three groups of students affect students' oral abilities. In order to do so, descriptive statistics and two-way analysis of variance were employed to find the differences of oral abilities resulting from these three task types and the three groups of students. Content analysis was also used to describe the differences in terms of accuracy, fluency, complexity, and comprehension.

3.2 Research Design

There are two kinds of variables investigated in this research. Three task types are task inputs and students' oral abilities scores of are task outcomes.

3.2.1 Task Inputs refer to three task types: answering questions, describing pictures, and transferring information. Three groups of students are advanced, intermediate, and beginning students.

3.2.2 Task Outcomes refer to the CMLSTT scores in terms of four aspects of language performance: accuracy, fluency, complexity, and comprehension.

3.3 Population

The population of this study is 253 first-year students whose majors are in English and Business English from the Faculty of Education and Faculty of Humanities and Social Sciences at Buriram Rajabhat University. They enrolled in the Communicative English I course in the first semester of the academic year 2009. These students all have similar language backgrounds with no experiences abroad and

no or very little exposure to English outside classes. They have a similar number of years in learning English and similar learning habits in practicing English outside classes. In addition, they have some basic skills in operating computers.

The sample consists of 84 students drawn by using stratified random sampling from the population in order to obtain three groups of students, namely, advanced, intermediate, and beginning. The 20 advanced students are the students whose scores from the Cambridge Key English Test (KET) are at or above the level of the 80^{th} percentile or whose scores are above or equal to 25 points. The 20 intermediate students are students whose scores are between 70^{th} and 40^{th} percentiles or whose scores are between 16-22 points. The 20 beginning students are the students whose scores are at or under the 25^{th} percentile or whose scores are below or equal to 14 points. Eight students from each group were randomly selected to be the sample for the pilot study, and 20 took part in the main study.

3.4 Research Instruments

There are two major types of research instruments employed in the study: a standardized English test and developed instruments.

3.4.1 The Standardized English Test

The Cambridge Key English Test (KET) was used as the standardized test to group the students according to their oral abilities. The difficulty of KET is at the A2 Level of the Common European Framework of Reference for Languages (CEFR). It is composed of three papers (Reading and Writing, Listening and Speaking). Since the researcher wished to classify the students based on their aural-oral abilities, only Listening and Speaking Papers were used.

Listening Paper : This sub-test takes 30 minutes and has five parts (twenty-five items total) to assess students' abilities to understand the factual information presented by the recorded materials. One mark is given for each correct answer; this paper carries 25 marks.

- Part 1: (Items 1-5) The students listen to five short conversations; there is one question is asked for each conversation. Three pictures are presented as the choices; the students can choose only one picture to best answer the question.

- Part 2 : (Items 6-10) The students listen to a short conversation and match 5 given words with the related information according to the content from the conversation.

- Part 3 : (Items 11-15)The students listen to a short conversation and choose the given answers for the 5 questions relating to the contents in the conversation.

- Part 4 : (Items 16-20) The students listen to a short conversation and write information from the conversation to complete the given form according to the guided words.

- Part 5 : (Items 17-20) The students listen to a short talk about a general topic and write the information to complete the given form with the information from the short talk according to the guided words.

Speaking Paper : This test takes up to 10 minutes and has two parts to assess the students' ability to interact in conversational English with an examiner and with another candidate during pair work. Students have to answer the questions about themselves and what they read in the card. This paper carries 25 marks. It consists of two parts.

- Part 1 : Students answer the examiner's questions about their personal information and daily life.

- Part 2 : Students work in pairs, taking turns asking the questions following the guided word cues and answering the questions by giving the information of people, places or lessons in the card. Students' speaking abilities are scored by the examiner based on the criteria of 3 language elements: grammar and vocabulary, pronunciation, and interactive communication.

3.4.2 Developed Instruments

In this study, two research instruments- the Computer-mediated Listening-Speaking Test Tasks (CMLSTT) and CMLSTT Rating scales were developed by the researcher.

1. Computer-mediated Listening-Speaking Test Tasks (CMLSTT)

The CMLSTT contains 3 tasks: answering questions, describing people, and transferring information. The time allowed to complete the test is about fifteen minutes. The test under the priori and posteriori validation processes is presented in Tables 3.1, 3.2, 3.3, and 3.4 respectively.

- Priori Validation Stage

In order to determine the constructs to be measured, the course description and the objectives of "Communicative English I" course, the constructs of "standards of Thai University English Foundation Courses I and II" and "the can-do statements" in listening and speaking skills of the level 1 of ALTE Framework, www.alte.org/cando/framework/level1.php) were explored. Table 3.1 shows the comparison of language functions as reviewed from different sources. The functions mentioned by two sources were selected to be included in the questionnaires. After that the constructs of the CMLSTT were selected from 75% of the answers to the questionnaires distributed to 14 English lecturers at BRU. The questionnaire is given in Appendix A. The speaking constructs and topics selected by the BRU lecturers are presented in Tables 3.2 - 3.4.

Table 3.1: Language functions based on the course description and other Thai and foreign sources.

Language Functions	Course[1]	CHE[2]	ALTE[3]	Nunan[4]
- Asking and giving personal information	/			/
- Giving personal information	/			
- Asking for information	/	/		/
- Talking about family and the chores	/			
- Talking about leisure activities	/	/		/
- Following the instructions	/		/	
- Describing the locations of things	/		/	
- Asking and giving directions	/			
- Describing actions	/		/	
- Ordering food and drinks	/		/	/
- Describing things and asking the prices	/			
- Describing occupations and activities	/			/
- Talking about daily schedule and future plans	/		/	
	/		/	
- Talking on the phone and leaving the message	/			/
	/			/
- Talking about daily routines				
- Describing things	/			/
- Describing people	/		/	
- Planning the vacation	/		/	
- Asking about the past events				
- Filling in the application form		/		
- Asking for basic services				
- Expressing and sharing feeling, opinions and ideas				

Notes :

1. content from the textbook ,"Journal : Listening and Speaking 1"
2. the standards of Thai University English Foundation Courses 1 and 2
3. The Association of Language Testers in Europe.
4. the graded activities from levels 1-3 for listening-speaking skills proposed by Nunan (2004)

Table 3.2 : The Language Constructs Selected by the BRU Lecturers

Language Constructs	No. of respondents (N=14)
1. Asking and giving personal information	**13**
2. Following simple instructions	2
3. Describing the locations of things	4
4. Describing people	7
5. Describing objects	1
6. Asking and giving simple directions	**11**
7. Ordering food and drinks	3
8. Describing goods in shops and asking the prices	3
9. Describing occupations	3
10. Talking about daily routine	**11**
11. Talking about past events	1
12. Talking about vacation plans	-
13. Talking on the phone	4
14. Asking for basic services	3
15. Expressing and sharing feelings and ideas	3

The top three language constructs chosen by the lecturers were asking and giving personal information, asking and giving simple directions, and talking about one's daily routine.

Table 3.3 Related Topics of Three Language Constructs Selected by BRU Lecturers

a. Asking and giving personal information

Language constructs /topics	*No. of respondents*
1. Asking and giving personal information	
a. names	**13**
b. age/birthday	9
c. marital status	4
d. hometowns	**11**
e. occupations	8
f. telephone numbers	9
g. addresses	6
h. families	**10**
i. friends	5
j. abilities	6
k. interests	**10**
l. leisure activities	7
m. others	-

Table 3.3 (cont.) Related Topics of Three Language Constructs Selected by BRU Lecturers

b. Asking and giving simple directions

Language constructs /topics	***No. of respondents***
6. Asking and giving simple directions	
a. a university	**10**
b. a city hall	4
c. a police station	**10**
d. a post office	9
e. a department store	7
f. a food store	6
g. a market	**10**
h. a bakery	3
i. a bookstore	5
j. a hospital	9
k. a bus terminal	**10**
l. a train station	9
m. others(*airport*)	1

c. Talking about one's daily routine

Language constructs /topics	***No. of respondents***
10. Talking about daily routine	
a. getting up	**10**
b. having meals	8
c. starting/ finishing studying	**10**
d. starting/ finishing working	**10**
e. transportation	8
f. leisure activities	8
g. going to bed	7
h. others (*household routine*)	1

Table 3.4 The Language Constructs and Related Topics Selected by the Majority of BRU Lectures

Language Constructs	Topics
1. Asking and giving personal information	- names - hometowns - families - interests
6. Asking and giving simple directions	- a university - a police station - a market - a bus terminal
10. Talking about daily routine	- getting up - starting/ finishing studying - starting/ finishing working

As seen in Tables 3.3 - 3.4, the top three topics related to the three language constructs chosen by the lecturers are as follows: The top three topics of the language construct, 'asking about and giving personal information' are names, hometowns, families and interests. The topics concerned with the language construct 'asking and giving simple directions' are a university, a police station, a market and a bus terminal. The topics of the language construct 'talking about one's daily routine' are get up, start/ finish studying and start/ finish working.

Since the construct "asking and giving personal information" garnered the most votes, it was used as a construct of the test. The construct "asking and giving simple directions" and "talking about one's daily routine" got the equal votes, thus "talking about one's daily routine" was randomly selected so that the test would not take too much time. Then, the researcher wrote the test specifications (See Appendix B) based on these two constructs.

The CMLSTT was developed in a series of the stages as follows :

- The CMLSTT was written according to the test specifications. The test was evaluated by five experts in the field of applied linguistics in terms of the content of

the test. (The details are in Appendix C.) Then the test validity is calculated by hand using the formula for the “the Item - Congruence Index”. The I.O. index value is 0.80.

$$\text{I.O.} = \frac{\sum X}{n}$$

$$= \frac{4}{5} = 0.80$$

- Revisions were made to improve the test according to the I.O. value and the experts’ suggestions and comments. Some of these are listed below:

In Task 1, the instruction "Give more details in each question." is not clear and should be changed to “Provide additional details to your answers” or “Give extra details with your answers.” Question no.5 in Task 1 should be more specific. The word “MC” in the instruction of Task 2 is difficult for students so the situation should be changed to be the MC to the students. In Task 4, question 2, “Where do you always have breakfast?”, the word “always” should be changed to “usually” since some students may not always eat at the same place and question 4, “What do you often do at the university?” is very vague so it should be changed to, “What university activities do you often do?”.

- Then, the test was put into a computer-based test format by a computer programmer by using the software program Microsoft Visual Studio 2005.

- Posteriori Validation Stage

In the pilot study, the CMLSTT containing six tasks was used. The study was carried out with 30 first-year students majoring in English from the Faculty of Education at BRU. They had similar characteristics with the sample of the main study This pilot group was excluded from the main study.

Due to the incomplete computer program in delivering the test via the internet, some problems occurred. The computers delayed in presenting the test so the students had to wait and they could not do the test at the same time. Another problem was that the students could not record their voices via the computer because of a problem in setting the audio modes. Complete answers from only 24 students could be analyzed.

Since the CMLSTT is an open-ended test (not multiple-choice format), the difficulty index and the discrimination index were analyzed by using the scores of the

students in the higher and lower groups or 25% of the sample (n=24) in the pilot study. Therefore, only the scores of the 6 responses in each group were calculated by hand using the following formulas (Suphat Sukamolsan, 2004: 50-54).

$$\text{IDiff (Difficulty Index)} = \frac{S_H + S_L - [(N_T)\,X_{min}]}{N_T\,(X_{max} - X_{min})}$$

S_H = the summation of scores of the examinees in the higher group

S_L = the summation of scores of the examinees in the lower group

N_T = the total number of N (N=12)

X_{max} = the highest score

X_{min} = the lowest score

$$\text{IDiff} = \frac{98 + 51 - [(12)(6)]}{12(19 - 6)} = \frac{77}{156} = 0.49$$

$$\text{IDisc (Discrimination Index)} = \frac{S_H - S_L}{N_H\,(X_{max} - X_{min})}$$

S_H = the summation of scores of the examinees in the higher group

S_L = the summation of scores of the examinees in the lower group

n_H = the number of the examinees in the higher group (N=6)

X_{max} = the highest score

X_{min} = the lowest score

$$\text{IDisc} = \frac{98 - 51}{6(19 - 6)} = \frac{47}{78} = 0.60$$

The scores of the students from each task were arranged in a rank order from high to low (the details are in Appendix D). The difficulty indexes of the CMLSTT is 0.50 and those of tasks 1-6 are 0.49, 0.45, 0.51, 0.57, 0.49 and 0.47 respectively

which means that the difficulty of all tasks are at the moderate level (between 0.20 – 0.80). The discrimination index of CMLSTT is 0.65, and those of the 6 tasks are 0.60, 0.68, 0.67, 0.62, 0.71 and 0.64 respectively. The discrimination indexes of all tasks are more than 0.30, and it means that all tasks can be used in the main study.

As seen in Appendix E, the mean scores of tasks 1-6 are 12.29, 11.58, 11.98, 13.38, 11.67 and 11.52 respectively and corresponding the standard deviations are 3.17, 3.73, 3.45, 2.98, 3.42 and 3.38. In order to select 3 tasks from the 6, the mean scores, standard deviations, discrimination and difficulty indexes were taken into account. Since the researcher wrote the parallel tasks in pairs (tasks 1 and 4, tasks 2 and 5, and tasks 3 and 6) only one task from each pair was chosen. Regarding the discrimination and difficulty indexes of each task, any task can be selected since there are no differences in these values. With regard to the mean scores, that of Task 1 when compared with Task 4 (12.29, 13.38), that of Task 2 when compared with Task5 (11.58, 11.67), and that of Task 3 when compared with Task 6 (11.98, 11.52) they are similar to each other. The value that helps decide is the standard deviation. Since this study needs a test which can classify the students according to their levels of language ability (namely, advanced, intermediate and beginning), the S.D. values of the tasks should be broad. Moreover, tasks 1-3 are based on the same language function (giving personal information) and tasks 4, 5, 6 rely on the same language function (talking about one's daily routine). Thus, only a set of tasks 1-3 (S.D. = 3.17, 3.73, 3.45) based on the language construct, "giving personal information", were selected to be used in the main study. Tasks 4-6 (S.D. = 2.98, 3.42 and 3.38) under the construct, "talking about one's daily routine" were eliminated.

In order to find test reliability, Cronbach's alpha reliability was computed by using the SPSS/PC Program Version 13.0. The reliability coefficient is .972 as presented in Appendix F.

Two teachers of English who have experienced in teaching a listening-speaking course for Thai undergraduate students at Rajabhat University for at least five years were invited to be the two raters. In Appendix G, the relationship between scores produced by two raters is presented in a scatter plot. Moreover, a correlation coefficient, Pearson Product Moment Correlation, was calculated to find the degree of relationship between two sets of scores or inter-rater reliability. Hatch & Farhady

(1982) suggested Pearson product-moment correlation since the scores of two raters are continuous and independent of each other and the relationship between the scores is linear. Also, the basic assumptions of Pearson product-moment correlation can be met. In the first round, the data shows the correlation value was .735, which is not high enough to ensure that grading from both raters is closely related. Thus, the discussion about the criteria and how to rate occurred afterwards to determine how to adjust the scores, and to assure that both raters understand the rating guidelines very well. The inter-rater reliability was calculated again. The correlation coefficient is .833 which is high enough to indicate that the two sets of scores go together well or that the scores from one rater are closely related to those from the other.

2. CMLSTT Rating scales

The CMLSTT rating scales were developed by the researcher based on

- the analytic descriptors of spoken language by the Council of Europe(2001)
- the oral proficiency test scoring categories by Brown (2001).

The rating scales included four language sub-skills: accuracy, fluency, complexity, and comprehension and each was divided into 5 score ranges: 0.0-1.0, 1.1-2.0, 2.1-3.0, 3.1-4.0, and 4.1-5.0. The scales, each of which had description of the different language proficiencies in grammar, vocabulary, pronunciation, cohesion and language function, were based on the grammatical knowledge in the model of language ability of Bachman and Palmer (1996). These scales were evaluated by two experts in the field of applied linguistics. The band descriptors were revised according to their suggestions (as seen in Appendix J).

3.5 Data Collection

The data collection was carried out in September 2009 in the Computer and Internet Center on the campus of BRU. In the process of data collecting, all computers were checked for readiness. Each computer was equipped with headphones and a microphone for recording the answers of the CMLSTT. The audio systems were appropriately set. Then, the three groups of the students were trained regarding how to manage the computer. When all the subjects were in their seats, the researcher explained the objectives and the significance of the study as well as the reasons for

using the CMLSTT. Both instruments were briefly described and the students were reassured that the scores would not affect their achievement grades or be used for any other purposes. Then they were assigned to do the three tasks of the CBSTT starting from tasks 1, 2 and 3, respectively. One group could be tested at a time due to the limited number of computers. Each task was presented in the same format- that is, the examinees would hear the instructions and then the visual materials or the questions would be presented. The test takers had the preparation time before performing their tasks and recording their answers to the questions within the given time before going on to the next item. The time allowed to complete the CMLSTT was about 15 minutes. Their responses were recorded onto the computers and retrieved via a flash drive. Later, the responses were assessed by two raters who have taught English for the tertiary levels for more than five years. The criteria are based on the analytical rating scales for accuracy, fluency, complexity and comprehension which have important and independent functions in oral performance and are normally used in the studies of task-based language testing (Skehan, 1998).

3.6 Data Analysis

To answer the first research question, “Do the advanced, intermediate and beginning students perform differently on three types of the CMLSTT”, the responses of the three groups of students performing the three types of CMLSTT were analyzed using descriptive statistics, i.e. means and standard deviations.

As for the second research question, “Do the different task types and groups of students affect the oral abilities of the students?”, two-way analysis of variance was used to compare the mean scores in four aspect of language performance of the three groups and those of the students performing the three task types. In order to locate the mean differences of the scores among the three groups and among those performing the three task types, the Scheffe′ Test was employed.

To answer research question 3, “How do the test takers’ oral abilities in each task type differ in accuracy, fluency, complexity, and comprehension?”, content analysis was carried out to analyze the content of the three groups’ responses in terms of four aspects of language performance: accuracy, fluency, complexity and

comprehension. The findings of the study and the discussions of the results are presented in Chapter IV.

Chapter IV

Results

This chapter presents the results of the study in accordance with its research questions. The results are presented in 3 parts:

Part 1: The oral abilities of the three groups of students performing the three types of the CMLSTT. Descriptive statistics were used to provide the answer for Research Question 1.

Part 2: The effects of 3 task types of CMLSTT and three groups of students on their oral abilities. Descriptive statistics and two-way ANOVA were employed to answer Research Question 2.

Part 3: The comparison of the oral abilities of the three groups of students in terms of accuracy, fluency, complexity, and comprehension in each task type. Content analysis was used to provide the additional answers for Research Question 3.

Part 1: The oral abilities of the three groups of students performing three types of the CMLSTT.

The scores of the three groups of students' oral abilities in performing the three task types are presented in Table 4.1. Means and standard deviations are included.

Table 4.1 Means and standard deviations of oral abilities of the three groups of students

Task Types	Groups of students	Oral Abilities							
		Accuracy		Fluency		Complexity		Comprehension	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Answering questions	Advanced	3.85	0.45	3.81	0.36	3.29	0.32	3.70	0.53
	Intermediate	3.30	0.76	3.17	0.80	2.70	0.65	2.88	0.77
	Beginning	2.42	0.89	2.52	0.94	2.25	0.80	2.26	0.93
	Total	**3.19**	**0.93**	**3.17**	**0.90**	**2.94**	**0.96**	**2.75**	**0.75**
Describing pictures	Advanced	3.44	0.52	3.58	0.72	2.89	0.64	2.57	0.69
	Intermediate	2.50	0.70	2.63	0.89	2.36	0.82	2.04	0.81
	Beginning	1.58	1.08	1.77	1.42	1.46	1.16	1.45	1.05
	Total	**2.50**	**1.10**	**2.66**	**1.27**	**2.02**	**0.97**	**2.24**	**1.07**
Transferring information	Advanced	3.34	0.73	3.62	0.61	2.96	0.56	3.34	0.75
	Intermediate	2.38	0.54	2.99	0.62	2.10	0.51	2.40	0.66
	Beginning	1.73	0.49	2.40	0.67	1.72	0.57	1.73	0.63
	Total	**2.48**	**0.89**	**3.00**	**0.80**	**2.49**	**0.94**	**2.26**	**0.75**
Total	Advanced	**3.54**	0.61	**3.67**	0.58	**3.05**	0.55	**3.20**	0.81
	Intermediate	**2.73**	0.78	**2.93**	0.80	**2.39**	0.70	**2.44**	0.81
	Beginning	**1.91**	0.92	**2.23**	1.09	**1.81**	0.93	**1.81**	0.94
	Total	**2.72**	**1.02**	**2.94**	**1.03**	**2.48**	**1.02**	**2.41**	**0.90**

Based on descriptive statistical analysis, it was found that of all the three groups, the advanced group performed the best in all three task types. Table 4.1 demonstrates that the advanced students performed the best in the tests of four aspects of language performance. In addition, the task of answering questions generated the highest scores in all four language skills. In regard to accuracy, advanced students had the higher scores (M=3.54) than the other two groups (M=2.73 and 1.91). Students

gained the highest accuracy scores in performing the answering questions task (M=3.19) when compared to the other two tasks (M=2.50 and 2.48).

As for fluency scores, compared with the intermediate and beginning students' scores, those of advanced students were the highest (M=3.67, 2.93 and 2.23). Students achieved the highest fluency scores from doing the answering questions task than the transferring information and describing pictures tasks (M= 3.17, 3.00 and 2.66).

Regarding complexity, among the three groups of students, the advanced students gained the highest complexity scores (M=3.05, 2.39 and 1.81). The students' complexity scores in performing the answering questions task (M=2.94) were the greatest when compared to those in the transferring information and describing pictures tasks (M= 2.49, 2.02).

With respect to comprehension, the comprehension scores of the advanced students were the highest (M=3.20) of the students (M=2.44, 1.81). The students gained the highest comprehension scores in doing the answering questions task when compared to their performance on the transferring information and describing pictures tasks (M=2.75, 2.26, 2.24). Since the standard deviations were quite low in all task types, the variability of the students' scores does not differ a lot.

Part 2: The effects of 3 task types of the CMLSTT and three groups of students on their oral abilities.

In order to determine if there were any significant differences in the students' oral abilities due to the three task types and groups of the students, the two-way ANOVA were employed on the four aspects of language performance: accuracy, fluency, complexity and comprehension. The results from the two-way ANOVA are presented in Table 4.2.

Table 4.2 The effects of the task types and groups of students on the accuracy scores

Source	Sum of Squares	df	Mean Square	*F*	Sig
task types	19.406	2	9.703	19.092	.000
groups	79.870	2	39.935	78.581	.000
task type*groups	1.623	4	.406	.799	.528
Error	86.903	171	.508		
Total	1523.598	180			
Corrected Total	187.802	179			

* $p < .05$

The results from the two-way ANOVA reveal that there was a significant main effect of task types on accuracy scores. This means that the students' accuracy scores on the three task types were different, F (2, 171) = 19.092, p ≤.001.

In terms of the main effect of groups of students, groups of students had a significant effect on the accuracy scores, (F = 78.581, df = 2,171, p ≤ .001). It indicates that the advanced, intermediate and beginning students achieved different accuracy scores.

As a significant difference was found in the accuracy scores influenced by the three task types, a Scheffe′ multiple comparison analysis was computed to identify which task type affects the scores in accuracy the most. The results are presented in Table 4.3.

Table 4.3 Results from the post-hoc comparison of the three task types, using accuracy as the dependent variable.

Multiple Comparisons

Dependent Variable: accuracy

Scheffe′

(I) Task types	(J) Task types	Mean Difference (I-J)	Std. Error	Sig.
Answering questions	Describing pictures	.6850*	.13015	.000
	Transferring information	.7075*	.13015	.000
Describing pictures	Answering questions	-.6850*	.13015	.000
	Transferring information	.0225	.13015	.985
Transferring information	Answering questions	-.7075*	.13015	.000
	Describing pictures	-.0225	.13015	.985

Based on observed means.

* The mean difference is significant at the .05 level.

The results from the post-hoc comparison indicate that the students obtained different accuracy scores in performing the three task types. The scores on answering questions were significantly different from those on the other two tasks. However, the accuracy scores on describing pictures were not significantly different from those on transferring information. To be more specific, students performing the task of answering questions got higher accuracy scores than those performing the other two task types (p ≤.001). The accuracy scores from the describing pictures task were a bit higher than those of the transferring information task.

In addition, in Table 4.2, there was also a significant effect of groups of students on the accuracy scores. A Scheffe′ post-hoc test was computed to identify how the accuracy scores of each group differed. The results from the Scheffe′ test are presented in Table 4.4.

Table 4.4 Results from the post-hoc comparison of the three groups of students, using accuracy as the dependent variable.

Multiple Comparisons

Dependent Variable: accuracy

Scheffe′

(I) group of students	(J) group of students	Mean Difference (I-J)	Std. Error	Sig.
Advanced	Intermediate	.8133*	.13015	.000
	Beginning	1.6317*	.13015	.000
Intermediate	Advanced	- .8133*	.13015	.000
	Beginning	.8183*	.13015	.000
Beginning	Advanced	-1.6317*	.13015	.000
	Intermediate	- .8183*	.13015	.000

Based on observed means.

***** The mean difference is significant at the .05 level.

The results from the post-hoc comparison show that the accuracy scores of the three groups of students are significantly different. The advanced students received the accuracy scores significantly different from the other two groups. The accuracy scores of the intermediate group were significantly different from those of the beginning group. Specifically, the advanced students got higher accuracy scores than the other two groups ($p \leq .001$). The accuracy scores of the intermediate group were also significantly higher than those of the beginning group.

The results from the two-way ANOVA (Table 4.2) show that there is no significant interaction effect between the task types and the groups of students since the *F*-ratio between these two variables is .799. The p-value, which is equal to .528 is greater than the .05 critical value. Therefore, the interaction between the three task types and the three groups of students did not affect the students' accuracy scores. This is shown in Figure 4.1

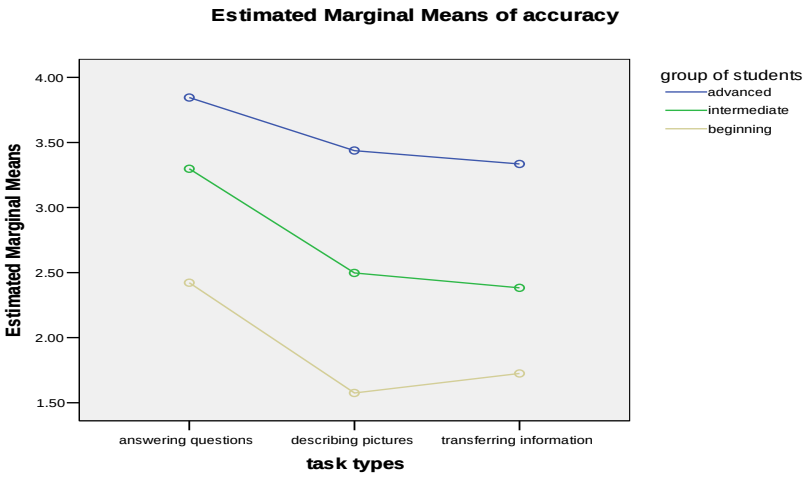


Figure 4.1 Plots of accuracy scores of the three groups of students

The plot in Figure 4.1 shows that there is no substantial crossing between the lines plotted from the main scores. Since each line of the profile plot follows the same pattern, that is, they are parallel, the interaction effect is not significant. The advanced and intermediate students gained higher accuracy scores in answering questions, describing pictures, and transferring information respectively. However, the order is different with the beginning students, who got higher accuracy scores in answering questions, transferring information and describing pictures. Table 4.5 illustrates the effects of the task types and groups of students on the fluency scores.

Table 4.5 The effects of the task types and groups of students on the fluency scores.

Source	Sum of Squares	df	Mean Square	*F*	Sig
task types	7.991	2	3.996	5.815	.004
group	62.294	2	31.147	45.333	.000
task types* groups	2.087	4	.522	.760	.553
Error	117.489	171	.687		
Total	1750.123	180			
Corrected Total	189.861	179			

* p <.05

The results from the two-way ANOVA reveal that the main effect of the task types is significant at .01. The *F* value for the task types is 5.815, which exceeds that of the critical value (df=2,171). It means that three different task types affect the fluency scores.

With regard to the main effect of the groups of students, the result shows that the groups of students significantly affect their fluency scores ($F = 45.333$, df = 2,171, $p \leq .001$). It indicates that the advanced, intermediate and beginning students achieved different fluency scores.

As significant differences were found in the fluency scores of the three task types to identify where these differences occurred, a multiple comparison analysis, a Scheffe′ Test was applied. The results are presented in Table 4.6.

Table 4.6 Results from the post-hoc comparison of the three task types using fluency as the dependent variable

Multiple Comparisons

Dependent Variable: fluency

Scheffe′

(I) Task types	(J) Task types	Mean Difference (I-J)	Std. Error	Sig.
Answering questions	Describing pictures	.5058*	.15134	.004
	Transferring information	.1642	.15134	.556
Describing pictures	Answering questions	-.5058*	.15134	.004
	Transferring information	-.3417	.15134	.081
Transferring information	Answering questions	-.1642	.15134	.556
	Describing pictures	. 3417	.15134	.081

Based on observed means.

* The mean difference is significant at the .05 level.

Table 4.6 demonstrates the results of the post-hoc comparison, where a significant difference was found in the fluency scores on two task types. The fluency scores from the answering questions task were significantly different from those on the describing pictures task. However, a significant difference was not found between the fluency scores from the answering questions and transferring information tasks as well as those from the describing pictures and transferring information tasks. Therefore, the students gained higher fluency scores on the answering questions task than on the describing pictures tasks. The students performing the task of answering questions receive a slightly higher accuracy scores than those performing the task on

transferring information. The fluency scores on the transferring information task were greater than those on the describing pictures task.

There is also a significant difference of the groups of students in regards to the fluency scores (Table 4.5). The Scheffe′ test was utilized to identify the score differences among the three groups of students. The results from the Scheffe′ test are presented in Table 4.7.

Table 4.7 Results from the post-hoc comparison of the three groups of students, using fluency scores as the dependent variable

Multiple Comparisons

Dependent Variable: fluency

Scheffe′

(I) group of students	(J) group of students	Mean Difference (I-J)	Std. Error	Sig.
Advanced	Intermediate	.7392*	.15134	.000
	Beginning	1.4408*	.15134	.000
Intermediate	Advanced	-.7392*	.15134	.000
	Beginning	.7017*	.15134	.000
Beginning	Advanced	-1.4408*	.15134	.000
	Intermediate	-.7017*	.15134	.000

Based on observed means.

***** The mean difference is significant at the .05 level.

The results of the Scheffe′ Test show that significant differences were found in the fluency scores of the three groups of students. The fluency scores of the advanced group were significantly different from those of the other two groups. There were also significant differences in the fluency scores of the intermediate and the beginning group. To be more specific, the advanced students performed more fluently than the other two groups did ($p \leq .001$). Similarly, the intermediate group obtained the higher fluency scores than the beginning group did.

Furthermore, Table 4.5 shows that at the .05 significant level, the task types and the groups of students had no significant interaction effect on fluency scores since the *F*-value was .760 (df 4,171), $p = .553$. Therefore, the interaction between the three task types and the three groups of students did not affect the students' fluency scores. This is shown graphically in Figure 4.2

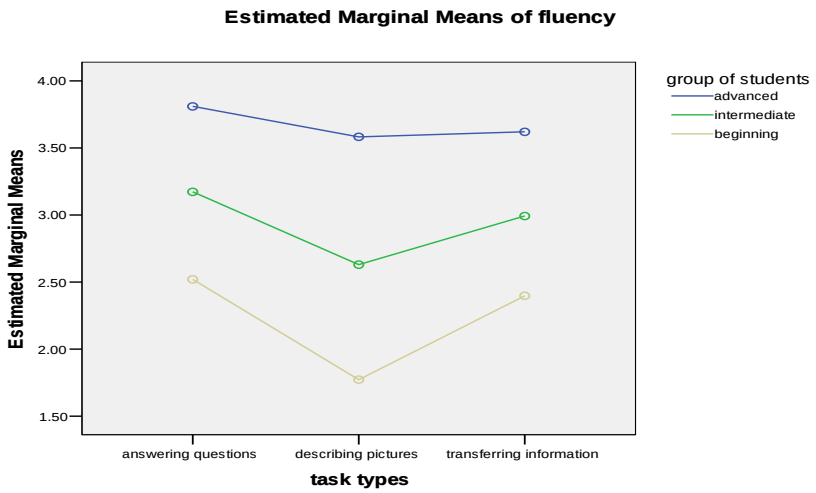


Figure 4.2 Plots of fluency scores of the three groups of students

As can be seen from Figure 4.2, the lines do not cross one another. The parallel lines indicate a lack of a significant interaction effect between the task types and the groups of students. In doing the three task types, the advanced group spoke the most fluently. Arranged from the highest to the lowest fluency scores, the orders of the tasks performed by the three groups of students were, namely, answering questions, transferring information and describing pictures.

Table 4.8 The effects of the task types and groups of students on the complexity scores

Tests of Between-Subjects Effects

Dependent Variable: complexity

Source	Sum of Squares	df	Mean Square	*F*	Sig
task types	10.079	2	5.039	10.063	0.000
groups	46.253	2	23.136	46.181	0.000
type * Groups	1.964	4	0.491	0.980	0.420
Error	85.633	171	0.501		
Total	1191.315	180			
Corrected Total	143.928	179			

* p <.05

Regarding complexity, the analysis of the two-way ANOVA, interestingly shows that there are significant differences in complexity scores due to the three task types (F = 10.063, df =2,171, p ≤.001). As shown in Table 4.8, a significant result was found. It shows that the students gained different complexity scores in performing the three task types.

With regards to the main effect of the groups of students, the two-way ANOVA in Table 4.8 reveals that a significant difference was found (F value = 46.181, df = 2,171, p ≤.001). It indicates that the complexity scores of the advanced, intermediate and beginning groups were significantly different.

Since there was a significant main effect from the three task types, a Scheffe′ multiple comparison analysis was computed to identify for which task type the students had the highest complexity scores. The results are presented in Table 4.9.

Table 4.9 Results from the post-hoc comparison of the three task types, using complexity scores as the dependent variable

Multiple Comparisons

Dependent Variable: complexity

Scheffe′

(I) Task types	(J) Task types	Mean Difference (I-J)	Std. Error	Sig.
Answering questions	Describing pictures	.5117*	.12920	.001
	Transferring information	.4917*	.12920	.001
Describing pictures	Answering questions	-.5117*	.12920	.001
	Transferring information	-.0200	.12920	.988
Transferring information	Answering questions	-.4917*	.12920	.001
	Describing pictures	. 0200	.12920	.988

Based on observed means.

* The mean difference is significant at the .05 level.

Table 4.9 shows the results of the post-hoc comparison, where a significant difference was found in the complexity scores of the three task types. The complexity scores on the answering questions task were significantly different from those on the describing pictures and transferring information tasks. However, a significant difference was not found between the complexity scores on the describing pictures and transferring information tasks. Therefore, the students gained higher complexity scores on the answering questions task than on the other two tasks.

The two-way ANOVA in Table 4.8 also reveals that a significant effect of the groups of students on the complexity scores was found. So, the Scheffe′ test was applied to check these significant differences. The results from the Scheffe′ test are presented in Table 4.10.

Table 4.10 Results from the post-hoc comparison of the three groups of students, using complexity scores as the dependent variable

Multiple Comparisons

Dependent Variable: complexity

Scheffe′

(I) group of students	(J) group of students	Mean Difference (I-J)	Std. Error	Sig.
Advanced	Intermediate	.6600*	.12920	.000
	Beginning	1.2408*	.12920	.000
Intermediate	Advanced	-.6600*	.12920	.000
	Beginning	.5808*	.12920	.000
Beginning	Advanced	-1.2408*	.12920	.000
	Intermediate	-.5808*	.12920	.000

Based on observed means.

***** The mean difference is significant at the .05 level.

The Scheffe′ post hoc test shows that the complexity scores of the three groups of students are significantly different. The advanced group scored differently in complexity from the other two groups. There are also significant differences in the complexity scores of the intermediate and the beginning group. Specifically, the advanced students gained higher complexity scores than the other two groups did ($p \leq .001$). Similarly, the intermediate group got higher complexity scores than the beginning group did.

Table 4.8 shows that the interaction between the task types and the groups of students on complexity scores is insignificant, ($F = .980$, df = 4,171, p ≥05). This indicates that three task types and the three groups of students did not affect the students' complexity scores significantly. Figure 4.3 shows the plots of the mean scores.

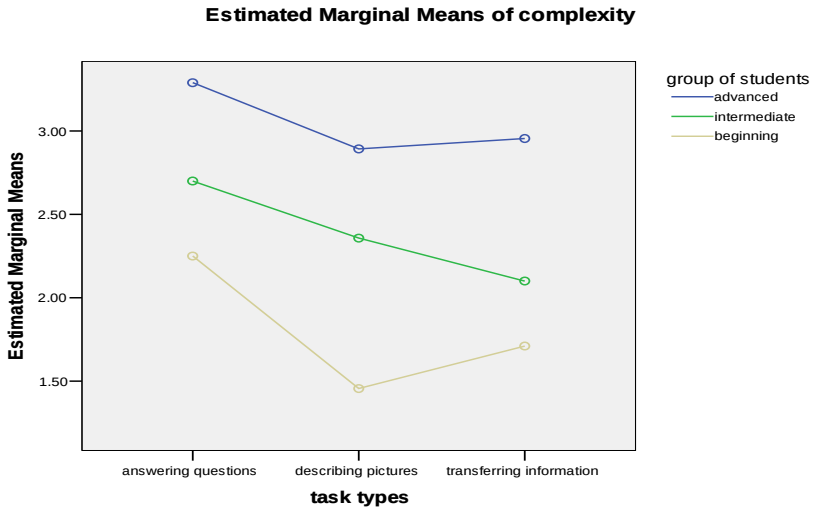


Figure 4.3 Plots of complexity scores of the three groups of students

Figure 4.3 shows that each line of the profile plot almost follows the same pattern**.** It reflects that there is no significant interaction effect between the task types and groups of students. The fact that the line for the advanced group is higher than those of the intermediate and the beginning groups shows that across all types of tasks, the advanced students had higher complexity scores than the other two groups. The advanced and the beginning groups gained higher complexity scores on answering questions, transferring information and describing pictures respectively. Although the intermediate students got the highest complexity scores on answering questions, they got better complexity scores on describing pictures than on transferring information.

Table 4.11 The effects of the task types and groups of students on comprehension scores

Tests of Between-Subjects Effects

Dependent Variable: comprehension

Source	Sum of Squares	df	Mean Square	*F*	Sig
task types	25.531	2	12.766	21.352	.000
group	58.226	2	29.113	48.695	.000
type * Group	1.425	4	0.356	0.596	.666
Error	102.236	171	0.598		
Total	1297.965	180			
Corrected Total	187.418	179			

* p <.05

The results from the two-way ANOVA in Table 4.11 show that there is a significant effect of the three task types on comprehension scores ($F = 21.352$, $df = 2,171$, $p \leq .001$). This indicates that students performing the three task types gained different comprehension scores.

In regards to the main effect of the groups of students, the two-way ANOVA in Table 4.11 also reveals that the groups of the students significantly affected their comprehension scores ($F = 48.695$, $df = 2,171$, $p \leq .001$). This indicates that the comprehension scores of the advanced, intermediate and beginning groups are significantly different.

Since a significant difference was found in the comprehension scores influenced by the three task types, a Scheffe′ multiple comparison analysis was applied to identify where the differences occurred. The results are presented in Table 4.12.

Table 4.12 Results from the post-hoc comparison of the three task types, using comprehension scores as the dependent variable

Multiple Comparisons

Dependent Variable: comprehension

Scheffe′

(I) Task types	(J) Task types	Mean Difference (I-J)	Std. Error	Sig.
Answering questions	Describing pictures	.9225*	.14117	.000
	Transferring information	.4558*	.14117	.006
Describing pictures	Answering questions	-.9225*	.14117	.000
	Transferring information	-.4667*	.14117	.005
Transferring information	Answering questions	-.4558*	.14117	.006
	Describing pictures	. 4667*	.14117	.005

Based on observed means.

* The mean difference is significant at the .05 level.

Table 4.12 demonstrates the results from the post-hoc comparison that there is a significant difference in the comprehension scores of the three task types. The comprehension scores on the answering questions were significantly different from those on the other two tasks. In addition, the comprehension scores on the describing pictures task were significantly different from those on the transferring information. Specifically, the students performing the task of answering questions gained higher

comprehension scores than those performing the other two task types ($p \leq .001$ and $p \leq .01$). The comprehension scores on the transferring information task were higher than those of the describing pictures task.

The two-way ANOVA in Table 4.11 also reveals that there is also a significant effect of the groups of the students on the comprehension scores. So, the Scheffe′ test was utilized to identify the score differences among the three groups of the students. The results from the Scheffe′ test are presented in Table 4.13.

Table 4.13 Results from the post-hoc comparison of the three groups of students, using comprehension scores as the dependent variable

Multiple Comparisons

Dependent Variable: comprehension

Scheffe′

(I) group of students	(J) group of students	Mean Difference (I-J)	Std. Error	Sig.
Advanced	Intermediate	.7650*	.14117	.000
	Beginning	1.3908*	.14117	.000
Intermediate	Advanced	-.7650*	.14117	.000
	Beginning	.6258*	.14117	.000
Beginning	Advanced	-1.3908*	.14117	.000
	Intermediate	-.6258*	.14117	.000

Based on observed means.

***** The mean difference is significant at the .05 level.

The Scheffe′ post hoc test reveals that the comprehension scores of the three groups of students are significantly different. The advanced students got higher comprehension scores than the other two groups. There are also significant differences in the comprehension scores of the intermediate and the beginning groups ($p \leq .001$). In other words, the comprehension scores of the advanced students were higher than those of the other two groups. Similarly, the comprehension scores of the intermediate group were also higher than those of the beginning group.

Furthermore, Table 4.11 shows that at the .05 significant level, the task types and the groups of students had no significant interaction effect on comprehension

scores,(F = .596, df = 4,171, p ≥ .05). Therefore, the interaction between the three task types and the three groups of students did not affect the students' comprehension scores significantly. This can be seen from the plot in Figure 4.4.

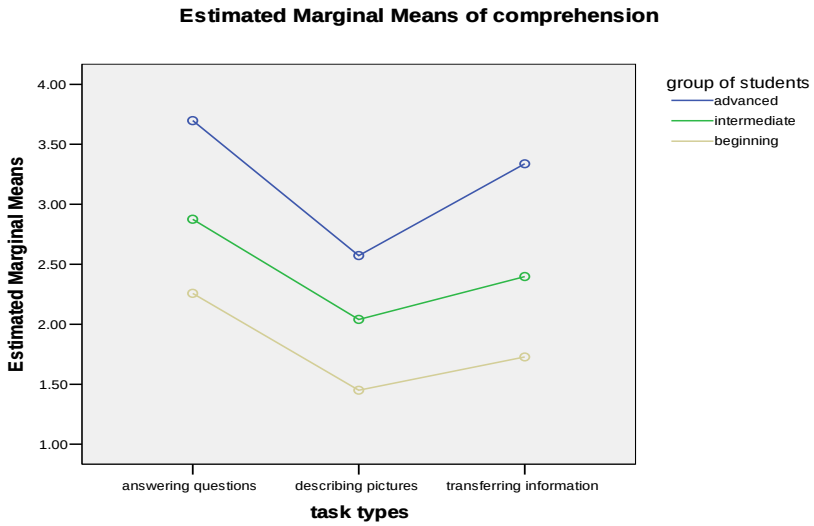


Figure 4.4 Plots of comprehension scores of three groups of students

In the profile plot of the data in Figure 4.4, the lines showing the means are parallel. This suggests the absence of a significant interaction effect between the task types and the groups of students on their comprehension scores. The mean line of the advanced group which is higher than those of the intermediate and the beginning groups means that the advanced group gained significantly higher comprehension scores than the other two groups. The tasks of answering questions, transferring information and describing pictures are arranged from the highest to the lowest respectively.

Part 3: The comparison of the oral abilities of the three groups of test-takers in terms of their accuracy, fluency, complexity, and comprehension in each task type.

Task 1 : Answering questions

Accuracy

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary Std 19 Most of the words were correct to convey the meaning intended. However, a word was translated from Thai word as 'play the internet' Std 21 All of the words used were correct. Std 2 The student thought of Thai word when he could not think of an English word as in the word 'play sport' but he could correct himself after that by changing it to 'play badminton'.	**1. Vocabulary** Std 44 The student was able to use a number of word choices to convey his message. Std 42 The student used all of the words correctly. Std 49 Most of the words are correct except the word "town" used instead of the Thai word, "Amphur". Std 46 A number of words were used to convey his meaning. Talking about what he likes to do in his free time, he used the sentence describing his daily routine instead as in 'I play basketball every day'.	**1. Vocabulary** Std 73 The student could think only of the word 'farmers' to tell what his parents do. Std 64 The student used isolated simple words to answer most of the questions. Std 76 The student was able to give a few words relevant to the questions. Std 62 A few common words about his birth date were used to answer the question.

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary (cont.) Std 16 The word 'business' and 'personal business' were used instead of 'businessman' or 'seller'. Moreover, the students said Thai words when he could not think of the English word as in the words 'tumbon, amphur and changwat' or translated from Thai verb like in 'use the internet'. Std 18 The student could not think of the word describing their parents' occupation so he used the English word that is close to Thai word; "do a farmer".	**1. Vocabulary (cont.)** Std 37 All of the word are correct and he could produce only a few words and sentences.	**1. Vocabulary (cont.)** Std 67 The student could think of only one word: 'brother' to answer the question, 'How many brothers and sisters do you have?'.

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar Std 19 The student shows the high degree of grammatical accuracy. Only a mistake was found in that he did not add 's' after the plural form of 'two sister'. Std 21 The student used a cardinal number instead of an ordinal number forl his birth date as in 'June eighteen'. He used wrong pronoun 'he' referring to his mother. Also in the sentence 'I like listen to music', he did not know that the gerund or to infinitive can follow the . verb "like"	**2. Grammar** Std 44 The student made basic grammatical mistakes. For example, the main verb in 'My birthday twenty-one March, nineteen eighty-eight.' was missing. He also put a quantifier in the wrong place instead of saying 'two sisters, he used 'sister two', which was clearly based on the form of Thai language. In addition, he expressed negation incorrectly; instead of saying 'I don't have any brother or I have no brother', he used 'none brother. He failed to pluralize nouns and use a possessive adjective where required, like 'I read book' or 'father and mother are farmers'. Std 42 The student produced a few simple grammatical errors; he made a mistake in the subject-verb agreement like in 'My parents	**2. Grammar** Std 73 All the sentences are grammatically incorrect. The student told the name with a phrase. In the sentence, 'My sister is one sister.', 'My sister' was used as a subject of the sentence, instead of 'I' and 'Is' was used as a main verb instead of 'have'. In addition, the student used the double subjects or wordiness as in 'Mother and father we are farmers'. Std 64 The student failed to produce his language in complete sentences, most of which have no subjects and main verbs. He also put a modifier in the wrong order and used the wrong comparative in 'young one' brother; it should be 'one younger' brother. The student used a cardinal number to tell his birth date instead of an ordinal number.

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) Std 2 The student used unnecessary conjunctions 'and' to connect some sentences that were not related to each other. He also misused the preposition 'on' instead of 'in' in front of the phrase 'my free time'. Std 16 To tell his birth date, the student used a cardinal number instead of an ordinal number like in 'twenty-seven of November'. An incorrect connector 'and' was used to link irrelevant sentences as in the sentences, 'My father is business, personal business and my mother is a housewife'. Lastly, he thought of a Thai word when he could not think of a correct form of English verb as in the sentence 'I no have' brother and sister'. In this	**2. Grammar (cont.)** is…'. He also failed to use a preposition 'in' in front of the phrase 'My free time', Parallelism was another mistake he made like in 'I like reading listen to music'. Finally, he did not use the definite article 'the' in front of 'family'. Std 49 There were no verbs in two sentences: 'My nickname Oh.' and 'Birthday three May one nine hundred and ninety.' Moreover, he did not know the way to tell the year he was born as "one nine hundred and ninety". In addition, he did not use ('s) to express possession in "my mother name" and in "my father name". Std 46 The student made a few grammatical mistakes and produced incomplete sentences. He used double subjects and double verbs, and he used	**2. Grammar (cont.)** Std 76 The student gave a series of nouns without using commas to divide each noun as in 'I have one brother mother father'. Also, he failed to use the preposition 'at' after the main verb 'live' to specify where his hometown is. Std 62 The student did not use the preposition 'on' in front of the date and month and add 's' to a plural noun like 'year'. He also used a cardinal number to tell his birth date instead of an ordinal number. In addition, he used a coordinator 'and' to connect his ideas without considering parallelism. Std 67 The student failed to create a complete sentence; he repeatedly gave the answer.

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) sentence he also neglected making the words plural forms and adding the quantifier 'any' before them. Std 18 The students used the sentences with incorrect forms of verbs as in the sentence "I born in the third of December." and also used the sentence without verb as in the sentence "My hometown in Surin." In addition, he did not know how to use the verb "like" which could be followed by gerund or to infinitive.	**2. Grammar (cont.)** the connector 'but' incorrectly like in 'My father, he is die but my mother is teacher.' Also, he did not use an indefinite article 'a' when telling an occupation, like in 'my mother is teacher.' Lastly, he failed to use the preposition 'in' in front of the phrase 'My free time,'. Std 37 There was no main verb in 'I born thirty May, one thousand ninety.' In addition, the student used the wordy subjects in 'Two people my father and my mother they are farmers.' Finally, he used a cardinal number instead of an ordinal number to tell his birth date, like in 'I born thirty May, one thousand ninety.'	

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
3. Pronunciation Std 19 The student used very natural and correct stress and intonation to convey the intended meaning. A few errors were found in /s/ ending sound as in 'year' and 'sister'. Std 21 The student can produce comprehensible speech with the appropriate use of stress and intonation. However, there was no /k/ ending sound in the word 'like'. Std 2 The speech was comprehensible with the correct use of stress and intonation. There was no /f/ ending sound in the word 'housewife'.	**3. Pronunciation** Std 44 The student produced constant rhythm, intonation and pronunciation but they did not cause unintelligibility. There were no ending sounds like /v/ in 'live', /k/ in 'book' and /m/ in 'game' etc. Std 42 The student tried to imitate the accent of the native speakers so he produced sentences with good stress and intonation but it seemed to be unnatural. Std 49 The student was intelligible but he pronounced the initial sound /θ/ incorrectly as in 'three' and no ending sound /v/ in 'have' and 'live'.	**3. Pronunciation** Std 73 The student produced a few short sentences with an acceptable rhythm of speech. Std 64 Some phrases were pronounced rather clearly. Std 76 It seemed like the student did not speak confidently and murmured some words. Std 62 Most of the words were pronounced without the correct ending sounds like /s/ in the word 'was', /n/ in 'born', /v/ in 'twelve' and /d/ in 'old'.

Task 1

Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
3. Pronunciation (cont.) Std 16 The sounds were sufficiently clear. He can correct himself when the mispronunciation occurred. Std 18 The student produced constant rhythm, stress and intonation so they did not attract the listener or convey the meaning.	**3. Pronunciation (cont.)** Std 46 The accent was comprehensible. He pronounced 'house' with no /z/ ending sound in and mispronunce the word 'live' as /laɪf/. Std 37 The accent was intelligible. He pronounced /t/ for the initial sound of the word 'thousand' and 'thirty'.	**3. Pronunciation (cont.)** Std 67 Only two phrases were pronounced with the incorrect /t/ sound in the word 'brother'.

The language abilities in accuracy of the three groups of students are concluded in the three aspects: vocabulary, grammar and pronunciation. In terms of vocabulary, most of the words the advanced group used were correct. Sometimes their words were influenced by their mother language or Thai in that they translated them from Thai as in these phrases 'play sport', 'use or play internet' or 'do a farmer'. Unlike the advanced group, the intermediate group could use some words correctly while the beginning one had a limited vocabulary so they could use only the simple ones.

With regard to the grammar ability, even the advanced group could not produce the utterances without the grammatical errors. The topics that the three groups always make the similar mistakes were the plural forms, preposition, using sentences without verbs, wordy subjects and verbs, unnecessary conjunctions, Thai sentence structure, an incorrect use of verb 'like' and using a cardinal number instead of an ordinal number. The differences in grammar abilities of the three groups were the advanced students could use the variety of structural forms while the beginning ones could produce only a few short and simple sentences.

Regarding the ability in pronunciation, the advanced group could produce comprehensible speech with the appropriate use of rhythm, stress and intonation to convey the meaning intended while the accent of the intermediate one was intelligible. The beginning students could produce acceptable rhythms of speeches. All groups shared the same pronunciation mistakes in that they pronounced the words without the ending sounds such as /s/, /k/,/f/,/v/ and with the incorrect initial sounds as /θ/.

Task 1

Fluency

Advanced Students	Intermediate Students	Beginning Students
Std 19 The student could express himself spontaneously with a natural flow. Std 21 The student was able to talk fluently without hesitations and pauses.	Std 44 The student spoke rather slowly with a few short pauses. Std 42 The student spoke rather fluently. However, a long pause occurred to interrupt his speech flow at the end of the speech.	Std 73 Due to a few short sentences with simple words, no fluent speech could occurred. Std 64 No fluent speech occurred because of some short phrases.

Task 1

Fluency (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 2 The utterances were produced rather fluently. A few hesitations sometimes occurred. Std 16 The student could talk intelligibly and smoothly with a few repetitions when searching for words. Std 18 He talked with constant levels of continuity and the speed rate was quite slow.	Std 49 Though his speech rate was rather slow, there were no hesitations and pauses to interrupt the flow of speech. Std 46 The student could produce the utterances with moderate speed rate. A few hesitations and word repetitions also occurred. Std 37 The student spoke rather slowly and sometimes said words by words so fluent speech cannot occurred.	Std 76 His speech was rather disfluent due to the inability to think of what to say. Std 62 Speech fluency could not be noticed due to an isolated sentence. Std 67 Speech fluency could not be found in two phrases.

In terms of the abilities in language fluency, the advanced students could speak fluently with minimal hesitations that did not interrupt the flow of speech. The speeches of the intermediate students were rather fluent due to a few pauses and hesitations. However, no fluent speech could be found in those of the beginning students since they could speak a few short sentences.

101

Task 1

Complexity

Advanced Students	Intermediate Students	Beginning Students
Std 19 Only simple sentences were used to answer each question. Only the conjunction 'and' was used to join words and sentences. Std 21 Though there were no connectors used in his speech, he gave more details by adding more sentences to modify the previous subjects. Moreover, he chose the adverb 'too' to modify the aforementioned verb as in the sentence, 'My mother is Lampueng. He is a farmer, too. Std 2 In spite of many conjunctions, 'and', they were used improperly to connect irrelevant sentences.	Std 44 The student used only the connector 'and' to link words or phrases, so there was no significant coherence. Std 42 The student produced isolated simple sentences with no connectors and cohesive devices. Std 49 Generally, the student produced isolated simple sentences. He used 'and' to connect information, but it seemed that the information was repeatedly included. Std 46 There is no coherence among sentences. Even if he used a connector 'but' to relate his ideas in 'My father, he is die but my mother is teacher.', it has nothing to do with the death of his father.	Std 73 There is no complexity in her statements; she created only simple sentences and phrases. Std 64 The student used only simple isolated words in simple structures and produced language with short utterances. Std 76 The student's response lacks complexity as there are no connectors and cohesive devices used to connect the ideas. Std 62 The student used only simple vocabulary in simple sentences with a connector 'and'.

Task 1

Complexity (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 16 A few conjunctions 'and' were used to join words and sentences. More sentences were used to add more details to the previous sentences. Std 18 The student joined some words and sentences only with the conjunction 'and'.	Std 37 The student produced a very low degree of coherence.	Std 67 The student used only one simple word, which completely lacked coherence and complexity.

As for the abilities in language complexity, most of the students in the advanced as well as the intermediate groups used simple sentences with the connectors 'and'. Instead of linking two sentences with the connectors, some students in these two groups added more sentences to modify the previous ones. Only simple words and sentences were used by the beginning group without any connectors.

103

Task 1

Comprehension

Advanced Students	Intermediate Students	Beginning Students
Std 19 The student could understand the instruction very well, including introducing himself. Std 21 The student could respond to all of the questions and add some more details to his answers such as his parents' names. Only one question that he missed was about his hometown. Std 02 His answers covered all that the questions need and he could add more details about his brother. Only an instruction that he missed is introducing himself.	Std 44 Overall, the student could include relevant information that he was asked, but sometimes he failed to express his complete thought in terms of meaning, like in 'I read book and listen and play a game'. Std 42 The student could understand some of the instruction and situation. Generally, he gave incomplete and irrelevant responses to complete the task. Std 49 The student showed that he could not catch up with some of the questions. The question about his parents' occupations, for example, he answered with their names instead. Also in the question about his hometown, he answered with where he lives now.	Std 73 The student was unable to catch all the asked questions. She failed to give all the answers correctly. She spelled her name even if it was not asked. Std 64 The student showed that he was able to understand what he was asked. However, she failed to answer one question about his parents' occupation. Std 76 The student was unable to catch all the asked questions, and he added irrelevant information to the question: How many brothers and sisters do you have?

104

Task 1

Comprehension (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 16 The student could understand the instruction and questions well. He also gave the additional details to describe her status as an only child. Std 18 The student could understand the instruction and questions quite well but did not add more information to each response.	Std 46 The student could understand most of the instruction and situation, but he failed to give one response to his birth date. Std 37 The student could understand most of the instruction and situation, but he failed to give one response to a number of brothers and sisters. Interestingly, he was able to put his orderly responses.	Std 62 The answers showed that the student was able to catch up with one question. But he gave an answer to the one that he was not asked. Std 67 The student was unable to catch up with most of the questions, only one of which was answered.

With respect to language abilities in comprehension, the responses of the advanced students showed that they could understand the instructions and situations well as they answered all of the questions and added some more details to their answers. Giving incomplete and irrelevant responses, the intermediate group could understand some questions while the beginning one could answer only a few questions.

Task 2: Describing pictures

Accuracy

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary Std 16 The student could not think of an English word so he tried to find the word that is close to Thai meaning as 'the first of Thailand' and 'the first of the world.' Std 20 The English word translated from Thai word was used like 'the one of Thai' to describe Paradon as a Thai professional tennis player. Std 18 The student could use the correct simple words to convey the meaning.	**1. Vocabulary** Std 35 The student could not think of related words to complete the task. However, he attempted to use the word 'champion' to describe Paradon's occupation, but it is not correct. Std 37 The student was not able to think of basic words to complete the task; it is obvious that he struggled to find an appropriate word to describe Paradon's occupation. Std 48 It's obvious that the student could not think of even a very basic word to tell Paradon's occupation, so he used a mere word 'tennis' instead of 'a tennis player'.	**1. Vocabulary** Std 84 The student failed to use a variety of vocabulary words, and he used a lot of incorrect words to express his intentions. For example, he used 'a badminton sport' twice to tell the occupation of Paradon instead of 'tennis player'. In addition, instead of using 'live', he used 'stay'. Std 86 The student used only a few words and a Thai word 'Okexercise' in response to the questions. Std 74 The student used a short simple sentence with a few words.

Task 2 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary (cont.) Std 27 The student could not think of the adverb 'since' so he used 'begin' to tell a particular time in the past as in 'He start playing tennis begin children.' Std 19 The student could use the correct simple word choices to convey the meaning at the elementary needs.	**1. Vocabulary (cont.)** Std 32 The student could not think of the words necessary for performing the task; he included irrelevant words, like 'table' or 'master', and in some parts of his response, he left it blank with cluelessness. Std 58 The student produced responses with basic words without a various number of vocabulary necessary for completing the task.	**1. Vocabulary (cont.)** Std 76 The student used short simple sentences with a few words. Std 79 The student used the repetition of words both tennis and sport in the sentence 'Paradon like tennis sport.'.

Task 2 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar Std 16 The student made basic grammatical mistakes. For instance, the subject-verb agreement, the student did not add 's' to the main verbs 'sit' when it was used with the third singular person 'he'. In addition, the present simple tense was used instead of present perfect tense in 'He play tennis since he was a child.' It does not make sense to use the adverb 'thus' with the sentences that are not cause and effect sentences like in 'He's married with Natalie and thus he sit on the motorcycle with Natalie.' Std 20 The incorrect pronoun 'she' was used for Paradon. The student could not use the correct form of tenses as in 'Paradon is playing tennis.'. Moreover, he did not know the appropriate form of verb such as	**2. Grammar** Std 35 The student could produce only three sentences with a few grammatical errors. Instead of saying 'Paradon Srichapan is a tennis champion in Thailand, he used 'Paradon Srichapan is champion of tennis of Thailand.' This is too wordy. In addition, there is no main verb 'is' and a preposition 'to'. Instead of 'he is married to Natalie Glebova.', he used 'And she married Natalie Glebova.' He also used a wrong reference. Std 37 The student could not construct complete sentences; he used prepositional phrases as subjects of the sentences. Those sentences are clearly based on the forms of Thai language; for example, 'In the family have five people.' and 'In picture has Paradon Srichapan.' Specifically, he had a problem in understanding the concepts of subject-verb	**2. Grammar** Std 84 The student was able to use complete sentences, but he made a lot of grammatical mistakes: subject-verb agreements, using adjectives as nouns, possessive adjectives, and prepositions. For the subject-verb agreement, the student did not add 's' to the main verbs 'live' and 'stay' even if he used the third singular person 'he'.The student also used adjectives as nouns like in 'My family is a Paradon have a happy' instead of 'His family is happy.' In addition, the student used incorrect pronouns 'my' instead of 'his'. Finally, he used double and incorrect prepositions, like in 'Paradon stay in at Natalie.' instead of 'Paradon lives with Natalie.' Std 86 The student made a grammatical mistake by using 'is' with 'play' as in Paradon is play tennis.

Task 2 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) the verb+s for the third singular subject as in 'She play tennis every day in tennis player.' or passive voice verb form as in 'He marry with Natalie Glebova.'. Std 18 The student could not use the correct form of tenses; i.e, he used 'have to' instead of present perfect tense to describe the action at the particular time in the past until now as in ' Paradon have to play tennis when he young and still play tennis now.' Passive voice did not used in the 'And after that Paradon is marries with Natalie.' He used pronoun instead of possessive adjective in 'She name Natalie.'.	**2. Grammar (cont.)** agreement; he used plural verbs even if he used singular subjects. Std 48 The student made a lot of grammatical mistakes compared to the produced content. He constructed very basic sentences with run-on and fragment and no main verbs. Std 32 The student committed syntactical errors in subject-verb agreement, double subjects and two main verbs in a sentence, run-on and fragment sentences. These incorrect qualities can be found in almost every sentence. Std 58 The student made a few common grammatical mistakes: double main verbs, misplaced main verbs, subject-verb agreement, and article.	**2. Grammar (cont.)** Std 74 The student used a short simple sentence correctly. Std 76 The student had a problem with a subject-verb agreement as he failed to put an 's' after the main verb 'play'. In addition, the main verb in the third sentence is missed. Std 79 The student made an error in a subject-verb agreement as he failed to put an 's' after the main verb of the third singular person 'like'.

Task 2 : Accuracy(cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) Std 27 Almost all of the sentences are grammatically incorrect. For example, he made a mistake in the subject-verb agreement like in 'Paradon he have one brother.' He also failed to add –s with the third singular person in 'He like…', 'he start…', 'he love…', 'he not work'. In addition, he did not used passive voice in 'He married Natalie'. Std 19 The student shows the high degree of grammatical accuracy. Only a mistake was found in that he misused the tense in 'He play tennis since he was young.'.	**2. Grammar (cont.)** Instead of 'he plays tennis', he used 'he is play tennis'. He also placed the main verb wrongly like in 'He married is Natalie.'; it should be 'He's married to Natalie'. In addition, he failed to use a singular verb after a singular subject. In addition, he used the definite article 'the' in front of a proper noun 'Thailand'.	

110

Task 2 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
3. Pronunciation Std 16 The student could talk naturally with the correct stress and intonation. Std 20 His rhythm, stress and intonation were acceptable. There was an error when he added the /s/ sound in an unnecessary place like 'hard**s** working'. Std 18 The articulations of the individual sounds were insufficiently clear so the listener had to pay attention to understand what he said. Std 27 There was strong influence from L1 in his rhythm and intonation but it did not cause unintelligibility. Some errors were found in using unclear ending sounds like/v/ in 'love', /k/ in 'like' and / ʃ / in 'much'.	**3. Pronunciation** Std 35 The student produced constant rhythm, intonation and pronunciation so the listener has to pay more attention to hear what he said. Std 37 The students' breaking groups of words and pronouncing without stress and intonation did not convey the meaning. Besides, there were no correct ending sounds like /v/ in 'love', /k/ in 'like', /s/ in 'tennis'. Std 48 Unnecessary /s/ sounds occurred in the words 'girlfriends' and 'Paradons'. The student pronounced without stress and intonation.	**3. Pronunciation** Std 84 The accent was noticeably influenced by L1 but it was intelligible. In addition, the student pronounced a vowel sound between the initial consonant cluster as in Thai like in the word '**sp**ort'/sʌpɔrt/ and '**st**ay' /sʌteɪ/. Std 86 The student did not pronounce the sentence clearly. Moreover, she said Thai word 'Okexercise' when she could not think of the English word, 'exercise'. Std 74 There were no errors in the pronunciation of the very short sentence.

Task 2 : Accuracy (cont.)

Advanced Students	*Intermediate Students*	*Beginning Students*
3. Pronunciation (cont.) Std 19 The student could produce comprehensible utterances with appropriate stress and intonation.	**3. Pronunciation (cont.)** Std 32 There was strong interference from L1 in rhythm, intonation and pronunciation. He also had problems in dividing the groups of words. Std 58 His accent was comprehensible . There was no /d/ ending sound as in the word 'married'.	**3. Pronunciation (cont.)** Std 76 The student pronounced language softly with constant rhythm. Std 79 The student spoke a sentence softly and unclearly .

The language abilities in accuracy of the three groups of students are concluded in the three aspects: vocabulary, grammar and pronunciation. In terms of vocabulary, the advanced group used simple vocabulary. Not only the students in the advanced group but the beginning students as well whose the English words which were influenced by Thai. While the advanced students used the English words that are close to the Thai language such as 'the first of Thailand/ the world', and 'the one of Thai', the beginning ones used Thai words when they could not think of English words like, 'Okexercise' instead of the verb 'exercise'. For the intermediate group they could not think of even the simple vocabulary due to their limited knowledge whereas the beginning one could produce only a few words in the simple sentences and sometimes used word repetitions like 'a tennis/badminton sport'.

As for the grammar ability, the advanced group could not speak the utterances without the grammatical errors. The topic that was the common mistake for all three groups was subject-verb agreement. The grammatical errors in tenses, adverb and passive voice for the advanced group differ from those of the other two groups. Also, the run-on sentences, double subjects and verbs, and articles were some problems for the intermediate one. It seemed like the beginning ones had a few errors due to their short and simple sentences.

With regard to the ability in pronunciation, the advanced group could talk naturally with the correct rhythm, stress and intonation. While the intermediate students produced constant rhythm of speech so their utterances conveyed the intentions at the elementary needs. All groups shared the same pronunciation problems in that their accents were influenced by L1 in pronouncing words and rhythms. Sometimes they added unnecessary/s/ behind the words. For the beginning group, they spoke softly and unclearly with a constant rhythm.

Task 2 : Fluency

Advanced Students	Intermediate Students	Beginning Students
Std 16 The student was able to use the language fluently with the minimal hesitations and pauses. Std 20 The utterances were produced rather slowly with a few hesitations.	Std 35 The student produced the language rather slowly with a long pause but he could keep on talking. Std 37 The student talked slowly with a few pauses and word repetitions.	Std 84 The students spoke rather slowly but the speech flowed smoothly without pauses and hesitation. Std 86 Fluency could not be noticeable due to only one sentence.

Task 2 : Fluency (cont).

Advanced Students	Intermediate Students	Beginning Students
Std 18 The student talked with slow levels of continuity and the speed rate. Std 27 The student could produce the utterances with moderate speed rate. Std 19 The student could produce the language fluently. His speech flew naturally and smoothly. A few hesitations occurred but not interfered the flow of speech.	Std 48 The student spoke rather slowly with a few pauses. Sometimes he said words by words and used the word repetition that obstructed language fluency. Std 32 His speech was interrupted by frequent /uh/ sounds, pauses and word repetitions so no language fluency could occur. Std 58 The student talked with normal levels of continuity.	Std 74 No fluent speech occurred because of a very short sentence. Std 76 Due to a few short sentences and some word repetitions, his speech was disfluent. Std 79 Speech fluency could not be found in one sentence.

With regard to the abilities in language fluency in performing Task 2, high, moderate and slow levels of the students' speech rates can be noticeable in the advanced group. While the intermediate students had moderate and slow flows of speech interrupted with long pauses, word repetitions and the /-uh-/ sound. However, no fluent speech could be found in those of the beginning students due to a few short sentences.

114

Task 2

Complexity

Advanced Students	Intermediate Students	Beginning Students
Std 16 The sentences were rather complex since the student tried to introduce many adverbs like 'since' to tell a particular time in the past and 'then' to describe the sequences of the events. However, using the adverb 'thus' with the sentences that are not causes and effects did not make any senses. Besides, the conjunction 'and' was also used to link the words. Std 20 Only the conjunction 'and' was used many times for the same purposes as in joining the adjectives. Std 18 The students extended the sentences by using the adverb 'when' and 'after that' to tell the sequences of time.	Std 35 The student could produce only a low-level response with a few sentences, using a very basic connector 'and'. Std 37 There are no cohesive expressions used in the speech, so there is no coherence among sentences. Std 48 There is no coherence among sentences as the student produced basic responses with isolated words and incomplete sentences. Std 32 The student produced basic responses with isolated words and incomplete sentences. There are no connectors used to link ideas, even basic ones.	Std 84 The student used isolated simple sentences without connectors and cohesive devices, which resulted in a lack of coherence among ideas. Std 86 The student produced a very low-level response with one sentence. Std 74 The student produced a very low-level response with one sentence. Std 76 The student produced a very low-level response with two sentences, which did not really answer the questions. Even if he tried to use the coordinator 'or, it was meaningless.

Task 2

Complexity (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 27 The student added more details by linking the sentences with the conjunction 'because' and 'and'. Std 19 The student used only the simple sentences. Only the adverb 'since' was found to link the sentences. More sentences were used to add more details to the previous sentences	Std 58 The student's response is not coherent as the student produced basic responses with basic sentences. He used a connector 'and' where it is not necessary	Std 79 The student used a very short utterance.

Regarding the abilities in language complexity, most of the students in the advanced as well as the intermediate groups used simple sentences with the connectors 'and'. Apart from linking two sentences with the connectors and adverb, an advanced student added more sentences to give more details to the earlier ones. The intermediate students used isolated and incomplete sentences or did not add any cohesive devices. Low-level responses with meaningless connectors were used by the beginning group.

116

Task 2

Comprehension

Advanced Students	Intermediate Students	Beginning Students
Std 16 The student could understand the instruction quite well except the last sentence which was irrelevant to Paradon's personal information as in 'He's married with Natalie and thus he sit on the motorcycle with Natalie.'. Besides, he may not have some background about Paradon's family so he could not give more details. Std 20 It seems like the student understood only some instructions since half of the responses were about Paradon's characteristics not his information. Std 18 His responses showed that he could understand what the task needs except in the sentence 'Paradon and Natalie sit on motorbike together.'	Std 35 Generally, the student failed to perform the task as he could not include all information necessary for completing the task. Std 37 It seems that the student could not understand the situation and instruction clearly, so he had no ideas to perform the task. He also fell shy to give more in-depth information. Std 48 The student did not fully understand the instruction and situation; he could produce only a small number of sentences while failing to include the very essence of information. Std 32 The student did not fully understand the instruction and situation; he could produce only a small number of sentences while	Std 84 The student quite understood the instruction and situation, but failed to give detailed information; he provided inadequate responses to the questions. Std 86 The student could not understand the instruction. It seems that he did not know how to perform the task. Std 74 The student could not understand the instruction. It seems that he did not know how to perform the task. Std 76 The student could not understand the instruction and relate it to the stories. He could not achieve the task.

Task 2 : Comprehension (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 27 The student could give most of the relevant answers with some additional details. Std 19 The student could give most of the answers the task needs but there were some information he missed such as Paradon's family and his interests.	failing to include the very essence of information. He also included irrelevant details in his responses. Std 58 The student could not understand the instruction and situation; he could produce only a small number of sentences while failing to include important information.	Std 79 The student may understood what he had to do but could not achieve the task because of the limited knowledge.

With respect to language abilities in comprehension, the advanced students could understand the instructions and situation quite well by giving the relevant answers. Giving incomplete and irrelevant details, the intermediate group could understand some instructions while the beginning one could give inadequate answers and some could not perform the task.

118

Task 3: Describing pictures

Accuracy

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary Std 21 The student could think of a sufficient number of basic words; he used the vocabulary provided in the task to complete it. Std 19 The student could think of a sufficient number of basic words. However, he could not think of an exact word to tell the occupation of Butree Puedpong's father; instead of telling an occupation, the student described where Butree Puedpong's parents work. Std 2 The student could think of a sufficient number of basic words provided in the task to complete the task.	**1. Vocabulary** Std 37 The student could think of a number of common words to produce his task. Std 38 The student could use a number of words related to the assigned task with incomplete sentences. However, he included an irrelevant word 'farmer' in the task. Std 47 The student could use a number of words related to the assigned task with incomplete sentences. Std 48 The student could use basic and common vocabulary to complete the task, but failed to include necessary words**.**	**1. Vocabulary** Std 66 The student used only simple vocabulary with incomplete sentences. Std 87 The student could not think of relevant words to complete the task. Std 79 The student could think of a small number of relevant words to complete the task. Std 74 The student could think of a small number of words to produce his task. Std 73 The student could think of a small number of common words to produce his task.

Task 3 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
1. Vocabulary (cont.) Std 1 A number of basic words provided in the task were used to complete the task. Std 20 The student used basic words provided in the task to complete the responses with isolated words.	**1. Vocabulary (cont.)** Std 44 The student could think of a number of basic words; normally he used the vocabulary provided in the task.	

120

Task 3 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar Std 21 The student made a few common grammatical mistakes: numbers and subject-verb agreement. To be elaborate, he used a cardinal number instead of an ordinal number to tell a birth date and did not put the definite article in front of it**,** like in 'Her name is Butree. She was born on sixteen of October, nineteen ninety.' In addition, he failed to use singular verbs, like in 'She live in Samutprakarn, Thailand.' and 'She like Hamtaro and Dan D2B.' Std 19 The student made several common grammatical mistakes: numbers, subject-verb agreement, and double subjects. He used a cardinal number instead of an ordinal number to tell a birth date and did	**2. Grammar** Std 37 Mostly, the student made a lot of grammatical mistakes. Firstly, he failed to use a possessive adjective correctly; instead of using 'Her name', he used 'She name'. Secondly, there are no main verbs in three sentences. Thirdly, he also made a mistake in the subject-verb agreement; he did not add an 's' to the main verbs even if he used the singular subjects. Finally, he used double subjects as in 'My father she name Chaowalit. My mother she name Suwanna.' Std 38 The student made a lot of grammatical mistakes. Most of the sentences are incomplete, having no main verbs and lacking subject-verb agreement.	**2. Grammar** Std 66 The student could not create sentences with correct structures; all of the sentences are grammatically incorrect. Std 87 The student made a few grammatical mistakes: main verbs and quantifiers. For example, he failed to use 'verb to be' when using adjectives as subject complements like in 'You beautiful and lovely.' Std 79 The student made a few mistakes in subject-verb agreement; he did not add an 's' to the main verb even if he used the singular subjects like in 'Butree like taekwando sport.' and 'She live in Thailand.'

Task 3 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) not put the definite article in front of it like in 'She was born on sixteen of October, nineteen ninety.' In addition, he failed to use verbs and subjects in agreement with each other. Finally, he used double subjects to create his sentence, 'Her father his name is Chaowalit.' Std 02 The student made several common grammatical mistakes: double subjects, numbers, subject-verb agreement, run-on sentences and articles. He used the double subjects as in 'The woman in this picture, her name is Butree.' Also, he used a cardinal number instead of an ordinal number to tell a birth date and did not put the definite article in front of it like in 'She was born on sixteen of October,	**2. Grammar (cont.)** Std 47 The student made a lot of grammatical mistakes. He created a lot of incomplete sentences with no main verbs and some lacking subject-verb agreement. Also, he used a cardinal number instead of an ordinal number to tell a birth date**.** Std 48 The student made a lot of grammatical mistakes. He created a lot of incomplete sentences with no main verbs. Also, he used a cardinal number instead of ordinal number to tell a birth date**.** Finally, he did not use an apostrophe's after nouns when expressing possessions, like in 'Butree father is Chaowalit. Mother name is Suwanna.'	**2. Grammar (cont.)** Std 74 The student made a few grammatical mistakes. He failed to use a possessive adjective correctly; instead of using 'Her name', he used 'She name'. There is also no main verb in the sentence 'Date of birth October sixteen nineteen ninety.' Std 73 The student made a few grammatical mistakes. He failed to use a possessive adjective correctly; instead of using 'Her name', he used 'She name'. he also made a mistake in the subject-verb agreement; he did not add an 's' to a main verb even if he used a singular subject, like in 'She live in Samutprakarn in Thailand.'

Task 3 : Accuracy(cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) nineteen ninety.' In addition, he failed to use verbs and subjects in agreement with each other. Finally, he did not use an indefinite article 'a' to tell an occupation in the sentence 'he is state enterprise officer.' Std 01 The student made several common grammatical mistakes: possessions, numbers, run-on sentences and references. He didn't use an apostrophe's when he expressed possession as in 'she name is Butree.' and 'She name is Suwanna.' Also, he used a cardinal number instead of an ordinal number to tell a birth date and did not put the definite article in front of it like in 'Her birthday is sixteen October, nineteen ninety.'	**2. Grammar (cont.)** Std 44 The student made a lot of grammatical mistakes. He created a lot of incomplete sentences with no main verbs. Also, he used a cardinal number instead of an ordinal number to tell a birth date**.**	

Task 3 : Accuracy(cont.)

Advanced Students	Intermediate Students	Beginning Students
2. Grammar (cont.) Std 20 The student produced his responses with a lot of common grammatical mistakes: fragments, possessions, numbers, run-on sentences and references.		

124

Task 3 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
3. Pronunciation Std 21 The speech was comprehensible with the appropriate use of stress and intonation. However, the /id/ ending sound in 'interested' was missing. Std 19 The accent sounds natural with the good stress and intonation to convey the meaning. Std 2 The accent was intelligible. However, a few errors in adding unnecessary /s/ sound were found behind the proper noun, Chaowalit while the necessary ones after the verb of the third singular subjects were missing.	**3. Pronunciation** Std 37 The stress and intonation were constant and they did not convey the meanings. Std 38 The student produced constant rhythm, intonation and pronunciation. Many errors in pronunciation were found in 'birthplace' as /plæs /, 'enterprise' as / plsɪ/, 'officer' as /ʃ/. Std 47 The student pronounced the words constantly without stress and intonation. He mispronounced the word 'state' as /steɪtɪ/ and sometimes the way he separated the words did not convey the meaning. However, he could self-correct his pronunciation.	**3. Pronunciation** Std 66 The student produced constant rhythm, intonation and pronunciation. In addition, he pronounced unclearly especially the ending sound like /v/in 'five', /θ/ in 'birthplace. He also mispronounced in the simple word like birthplace / plɪs/ Std 87 His accent is intelligible with good stress and intonation. Some ending sound like /v/ in 'live' or /t/ sound in 'sport' were missing. Std 79 The students speak softly and unclearly. Std 74 His accent was acceptable.

Task 3 : Accuracy (cont.)

Advanced Students	Intermediate Students	Beginning Students
3. Pronunciation (cont.) Std 01 The student could produce easily comprehensible speech. He pronounced a vowel sound between the initial consonant cluster as in Thai language like in the word '**st**ate' /sʌteɪt/. Std 20 The student was intelligible but some errors were found in the ending sounds like /v/ in 'live' as well as /θ/ in 'birth'. In addition, /s/ sounds were missing after verbs of the third singular subjects.	**3. Pronunciation (cont.)** Std 48 The accent was acceptable but sometimes adds unnecessary /s/ after 'sixteen' and 'Chaowalit'. Std 44 The student produced the utterances without stress and intonation. The proper names were mispronounced.	**3. Pronunciation (cont.)** Std 73 Constant rhythm, intonation and pronunciation sometimes caused unintelligibility.

The language abilities in accuracy of the three groups of students are concluded in the three aspects: vocabulary, grammar and pronunciation. In terms of vocabulary, the advanced as well as the intermediate group used basic words. Apart from using only simple words, the beginning one also gave the irrelevant words in the simple sentences.

With regard to the grammar ability, the students even the advanced group use the languages with lots of grammatical mistakes. The common errors for all three groups were the subject-verb agreement, cardinal numbers, possessions and no main verbs. The run-on sentences and double subjects were the problems of the intermediate one.

With respect to the ability in pronunciation, the advanced group could talk naturally with the correct rhythm, stress and intonation while the intermediate students produced constant rhythm of speech so their utterances could not convey the meaning intended. The pronunciation problems for the advanced group were ending sounds and pronouncing with unnecessary /s/. Mispronunciation was a problem of the intermediate group. For the beginning group, they spoke softly and unclearly with constant rhythm but some can be intelligible.

Task 3 : Fluency

Advanced Students	Intermediate Students	Beginning Students
Std 21 The student could talk fluently without hesitations and pauses. Std 19 The student could express himself spontaneously with a natural flow but sometimes the speech was interrupted by	Std 37 The speed rate was very slow especially when he pronounced the proper noun and also was interrupted by the word repetition. Std 38 The student spoke at the moderate speech rate.	Std 66 There were disjointed phrases instead of sentences in his utterances so he did not produce smooth flowing. Std 87 A flow of speech was at the moderate level with a few pauses.

Task 3 : Fluency (cont.)

Advanced Students	Intermediate Students	Beginning Students
hesitations. Std 2 The student produced the sentences with the normal speed rate. A long pause was occurred when he tried to say the house number. Std 1 The student could talk spontaneously and fluently. Std 20 The speech rate is moderate but sometimes slow when finding the suitable words.	Std 47 The student produced the utterances slowly with a few short pauses when he tried to pronounce some proper names. Std 48 The student produced the sentences at the moderate rate but have a few short pauses. Std 44 The student produced the language rather slowly without hesitations and pauses.	Std 79 Due to a few short sentences and with a few long pauses, they interrupted the speech fluency. Std 74 Speech fluency could not be found in two short sentences. Std 73 No fluent speech could be noticed.

With regard to the abilities in language fluency in performing task 3, high and normal levels of the students' speech rates with a few pauses could be noticeable in advanced group. The intermediate students had moderate and slow flow of speech interrupted with short pauses and word repetition. The beginning students produced language at the moderate and slow levels of speech. However, no fluent speech could be found in those of the beginning students due to a few short sentences.

Task 3

Complexity

Advanced Students	Intermediate Students	Beginning Students
Std 21 The student did not use any connectors to link the ideas among sentences; he produced relevant and complete responses . Std 19 The student used no connectors to connect ideas among sentences; he produced relevant and complete responses . Std 02 The student used no connectors, even basic ones, to connect ideas among sentences; he produced relevant and complete responses .	Std 37 There are no connectors used to connect ideas; the student produced relevant responses with simple sentences. Std 38 There are no connectors used to connect ideas; the student produced relevant responses with incomplete sentences with isolated words and phrases. Std 47 There are no connectors used to connect ideas; the student produced relevant responses with incomplete sentences with isolated words and phrases.	Std 66 There is no coherence among sentences as the student produced language with isolated words. Std 87 There is no coherence among sentences as the student produced language with simple sentences. Std 79 The student used only simple words with basic sentence structure lacking coherence. Std 74 The student used only simple words with basic sentence structure lacking coherence.

Task 3

Complexity (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 01 The student did not use any connectors, even basic ones, to connect ideas among sentences; he produced relevant and complete responses. Std 20 The student's responses showed no complexity and coherence as no basic connectors were used. The student relied on isolated words and short utterances to complete the task.	Std 48 There are no connectors used to connect ideas among sentences; the student produced relevant responses with incomplete sentences with isolated words and phrases. Std 44 There are no connectors used to connect ideas among sentences; the student produced relevant responses with fragment sentences with isolated words and phrases.	Std 73 The student produced only low-level responses with basic sentence structures.

In terms of the abilities in language complexity, most of the students in the advanced and the intermediate groups used simple sentences with no connectors and some intermediate students used isolated words in incomplete sentences to give the information. Simple sentence structures were used by the beginning group without any connectors.

Task 3

Comprehension

Advanced Students	Intermediate Students	Beginning Students
Std 21 The student could understand the instruction and situation by performing the function clearly and successfully and including all the necessary and relevant information provided in the task and needed to complete it. Std 19 The student could understand the instruction and situation by performing the function clearly and including all the necessary and relevant information provided in the task to complete it. Std 02 The student could understand the instruction and situation by performing the function clearly and including all the necessary and relevant information provided in the task to complete it.	Std 37 The student was successful to complete the task as he was able to transfer information to complete and relevant speech. However, the student was confused with his statement as he answered the questions using the given information as her own. Std 38 It seems that the student could understand the instruction and situation. However, he fell shy to transfer the given information to complete the sentences, and he included irrelevant information. Std 47 It seems that the student could understand the situation. He was successful to complete the task as he was able to transfer information to complete the task, but with incomplete decent sentences. However, the student was confused	Std 66 Obviously, the student could not understand the instruction and situation; he was unable to transfer information with sentences to complete the task. Std 87 It's clear that the student did not understand the instruction and situation. He produced incorrect and irrelevant responses; he used 'you' as the subjects of the sentences. Std 79 The student did not understand the instruction and situation; he could produce only a small number of sentences while the most essence of the task is minimal.

Task 3 : Comprehension (cont.)

Advanced Students	Intermediate Students	Beginning Students
Std 01 The student could understand the instruction and situation by performing the function and transferring information clearly, but some necessary and relevant information provided in the task was not included. Std 20 The student could somehow understand the instruction and situation, falling shy to transfer the information.	with his statement as he involved himself by using the possessive adjective 'My'. Std 48 It shows that the student did not catch up with some of the necessary information. He failed to transfer information to complete the task; most of the necessary information was missed. Std 44 It seems that the student could remember the information provided in the task, but he failed to transfer it with complete sentences.	Std 74 The student did not understand the instruction and situation; he could produce only two sentences, being unable to transfer the information to complete the task. Std 73 The student did not understand the instruction and situation; he could produce only two basic sentences but was unable to transfer information to complete the task.

With respect to language abilities in comprehension, the responses of the advanced students could understand the instructions and situation well by giving the relevant answers. Giving incomplete and irrelevant details, the intermediate group could understand some instructions while the beginning one could give inadequate answers and some could not perform the task.

Chapter V

SUMMARY, DISCUSSION, AND RECOMMENDATIONS

This chapter comprises the summary of the study, discussion of the results, the implications yielded by the study, and recommendations for teachers and for further studies.

5.1 Summary of the study

5.1.1 Purposes of the Study

There were three purposes in this study. The first one was to compare the oral abilities of the advanced, intermediate, and beginning students in terms of different listening-speaking task types, second, to investigate the effects of the three task types of the CMLSTT and three groups of students on the oral abilities, and lastly, to study the differences in the accuracy, fluency, complexity, and comprehension of the test-takers' oral performances in each task type.

5.1.2 The Sample of the Study

The sample consisted of 84 first-year students of BRU enrolling in Communicative English I course in the first semester of the academic year 2009. To obtain the students with three levels of oral abilities: advanced, intermediate and beginning, drawn using stratified random sampling from the population at the 80^{th}, 70^{th} - 40^{th} and 25^{th} percentile ranks of The Cambridge Key English Test (KET) scores.

5.1.3 Research Instruments

The instruments used for collecting the data in the study were the Computer - mediated Listening-Speaking Test Tasks (CMLSTT), The Cambridge Key English Test (KET), and CMLSTT rating scales.

The Computer-mediated Listening-Speaking Test Task (CMLSTT):The test was a semi-direct computer-mediated listening-speaking test developed by the researcher. The CMLSTT contains 3 tasks: answering questions, describing pictures, and transferring information. The students listen to prompts and questions from a video and tasks are delivered using the computer. The language function of the test is

‘giving personal information’, taken from the course description and objectives of the “Communicative English I” course .

The Cambridge Key English Test (KET): the test was used as a standardized test to classify the students into (1) advanced, (2) intermediate, and (3) beginning groups according to their oral abilities. Only two papers (Listening and Speaking) were used.

The CMLSTT Rating Scales : The scales were modified from the analytic descriptors of spoken language by the Council of Europe (2001), the oral proficiency test scoring categories by Brown (2001), and the grammatical knowledge in the model of language ability by Bachman and Palmer (1996). It covered four dimensions of language performance: accuracy, fluency, complexity, and comprehension.

5.1.4 Research Procedure

The three groups of the students were trained to operate the computer; they were told the objectives and the significance of the study as well as the reasons for using the CMLSTT. Next, they were assigned to do the three tasks of the CMLSTT presented from tasks 1, 2, and 3, respectively. The test takers recorded their answers to the questions within the given time, before going on to the next item. The responses, as recorded on the computers, were assessed later by two raters by using the CMLSTT analytical rating scales for accuracy, fluency, complexity, and comprehension.

5.1.5 Data Analysis

To answer the first research question, descriptive statistics was employed to determine whether there were any differences in the oral abilities of the three groups of students performing the three types of CMLSTT. For research question 2, the two-way analysis of variance was used to investigate the effect of the three task types and three groups of students on the oral abilities. The Scheffe′ Test was employed to identify the significant differences among the three groups and the three task types. With regard to research question 3, content analysis was used to qualitatively analyze the differences of three groups of students in accuracy, complexity, fluency, and comprehension scores.

5.1.6 Summary of Research Findings

The findings of the study in response to the research questions can be summarized as follows:

Of all the three groups of students, the advanced students performed the best and gained the highest scores in the four aspects of language performance: accuracy, fluency, complexity, and comprehension. In addition, the answering questions task generated higher scores of all.

There were significant main effects of task types on all aspects of language performance: accuracy, fluency, complexity, and comprehension. The students in the three ability groups scored differently in four aspects of language performance when performing the three task types.

In terms of accuracy, complexity and comprehension, the scores when answering questions were significantly higher than those for the describing pictures task and the transferring information task. The comprehension scores on the transferring information task were significantly higher than those on the describing pictures task. Regarding fluency, the students gained higher fluency scores on the answering questions task than on the describing pictures task.

In terms of the main effect of groups of students, the three groups of students had a significant effect on the accuracy, fluency, complexity, and comprehension scores. The advanced students got higher scores in the four dimensions of language performance than the intermediate and beginning students. In addition, the scores in the four aspects of language performance of the intermediate students were higher than those of the beginning group.

However, the interaction between the three task types and the three groups of students did not affect the students' oral abilities. The plots of four aspects of language performance of the students in the three groups showed both similarities and differences in the orders of three task types. The fluency and comprehension scores of the three groups of students arranged from the highest to the lowest were from the answering questions, transferring information and describing pictures tasks.

The order of the accuracy scores of the beginning group and that of the complexity scores of the advanced and the beginning group were the same as those mentioned previously. The advanced and intermediate students gained the highest to the lowest accuracy scores in the answering questions, describing pictures and

transferring information tasks, which is similar to the order of the tasks arranged from the highest to the lowest complexity scores of the intermediate students.

5.2 Discussions

The results of research questions one and two were that there were significant differences in the main effects of task types on the learners' oral abilities as seen in the accuracy, fluency, complexity and comprehension scores obtained from the answering questions task. These scores were significantly higher than those from the describing pictures and transferring information tasks.

Accuracy scores in performing the answering questions task were greater than the other two tasks. This finding corresponds with Skehan and Foster's study (1999) in that they concluded that the main interaction effect between task types and test conditions produced higher accuracy scores. The structured task under 'the watch then tell condition' generated the greatest accuracy. The clearer inherent sequential structure in the structured task together with the planned condition may have helped the learners gain better accuracy scores. It can be reasoned that the sequential structure of the task leads the test takers to follow the formulation of the task and then they can concentrate only on sentence forms. Planning time also eases the processing burden of the task. Additionally, the planning time available for the test takers might have allowed the students to plan their speeches or ease the cognitive loads by letting them focus only on the global task (Ortega, 1999). Similar to Ortega, increasing planning time or emphasizing accuracy attendance enhances accuracy (Hinkel, 2005). According to Mehnert (1998), accuracy increases significantly on exposition tasks with even one minute of planning time. Skehan (1998) also adds that many educators place an emphasis on the pre-task activity or the planning before the students engage in the learning activity, which lets the students to be more ambitions in not only organizing their ideas but also retrieving the supporting information. In the current study, though planning time was not studied, the test administration allowed the students to have the preparation time before responding to each task. It may let the students to have more time preparing what to say. In addition, the questions provided by the task may be used as the guidelines of their answers so they have not to create their sentence structures on their own. The chances to make errors may be less than

the task, which the students have to create their own sentence structures. Consequently, the higher accuracy scores were found in the answering questions task.

In addition, the answering questions tasks generated a more fluent performance than the other two tasks. It might be because the clearer sequential structure of the answering questions task, as a structured task, leads the students to access the macrostructure of the task. Then, the test takers do not need to spend much time thinking of what to say next or engaging in 'mid-task repair' so their speech flows are not interrupted (Skehan and Foster, 1999). The findings of this study supported the work of Teng (2007) in that the answering question task generated higher fluency scores. Since this task seems to be the most common task among EFL speaking tests and activities in Taiwan, the subjects likely had more opportunities to answer questions in English than to describe pictures or present the information, so they could do well with this kind of test. Like Taiwanese students, Thai learners are used to learning English through practicing the asking and answering of questions in class. It might be possible that they also gain the better fluency scores in this kind of test. Moreover, the familiar topics dealing with the students' personal information such as their dates of birth, birthplaces or families in this test task might have led to higher scores when compared to those in less familiar tasks.

The students gained higher complexity scores on the answering questions task than the other two tasks (i.e. describing pictures and transferring information). As in Skehan and Foster's study (1999), the complexity measures were affected by processing condition. Students performing the narrative tasks under the "watch then tell condition" as a "non-simultaneous task condition" in this study gained higher complexity scores. On the contrary, the simultaneous condition or keeping up with the video in real time could generate lower complexity. Thus, the non-simultaneous task condition helps the students to gain higher complexity scores than those performing in the simultaneous tasks. The findings of Skehan and Foster (1999) were supported by the results of Kawauchi (2005) in that "the planning time condition helps learners to access their maximum levels of lexical and structural knowledge which will enable them to use more complex language." Apart from the effect of processing conditions on language complexity, Foster and Skehan (1996); Skehan and Foster (1997) also found an influence of task types on complex languages. The 'personal task' and 'the decision task' as less-structured tasks in the second study in 1997 led to more

complex and differentiated outcomes than the structured ones in the first work, since they need more elaboration and justification in expressing their ideas. Another finding which the explanation contrasts with is that of the previous study is that of Teng (2007). Teng found that there were significant differences in the complexity because of different task types. Based on the idea of Skehan and Foster (1999) that processing loads affect language complexity, Teng concluded that the answering questions task with its structured format required a greater processing load than the picture description task, which has a more flexible format. Students were expected to provide answers directly to the questions so the task requires a greater processing load than the picture description task, which needs fewer processing loads to generate flexible descriptions. From the results of the current study, though the answering questions task as a structured task could create higher language complexity, it does not mean that the answering questions task in this study needs a greater processing load than the other two tasks. On the contrary, the students needed a greater processing load since they had to find the words and language structure by themselves in order to create the sentences to describe the given pictures, as well as to make the sentences based on the given words in the transferring information task. That the complexity scores were not as high as expected in these two tasks may not be due to the nature of the task, but to the language incompetence of the test takers. The students might not be able to think of their own words or structures to describe the given pictures. In addition, they may not have been able to create the sentences from the given words in the transferring information task.

‘Communication scores’ in performing the answering questions task were also found to be higher than the other two tasks. In the present research, giving related responses might not indicate that all of the participants could understand the instructions and situations of the answering questions task better than those of the other two tasks. It might be that they were more familiar with this kind of test since it has been regularly used in their English classes. Though the other two tasks were also used as language activities in this course, this might not help them understand how to respond correctly to the tasks, especially to the transferring task. In addition, misunderstandings might result from the inefficient listening skills, so they may not have caught what the task needed them to do. Therefore, some responses to the task were in the form of isolated words or phases. Similarly, the students did not get

accustomed to the describing picture task and sometimes could not keep up with the instructions and situations proposed in the task. Thus, irrelevant answers were given e.g. his characteristics, feelings, even the students' point of view toward the sport athlete in this task.

5.3 Implications

The CMLSTT is not intended to be used as an achievement test for any oral courses. It is designed to be a progress test for the specific language function to assess students' oral abilities. In addition, the scores obtained from the test can be used with those from other kinds of testing to make decisions on students' oral performances. As per the results of the study, the students gained more accurate scores when they performed the structured tasks under the planning time condition. In addition, the structured tasks also generated more language fluency. Furthermore, the learners could gain more complexity scores when they performed less structured tasks and when they were under the planning time condition. The selective task types and task conditions can be used to design instructional sequences with the balanced development of all four aspects of language performance. The language activities might start from the more controlled task under the planning time condition to the less structured task without the planning time. The information from content analysis, especially the accuracy part, that should be taken into account is about the students' errors in grammar, vocabulary and pronunciation. These details might be used as the content and the materials in the preparation course to enhance the first-year students' language performance before the semester begins or they might be used for preparing the remedial materials of each lesson.

5.4 Recommendations

5.4.1 Recommendations for teachers

Since the design of the CMLSTT was based on 'Communicative English I Course', the results of this study will shed light on the enhancement of the syllabus, in order to design appropriate course materials and teaching methods for this course or any oral courses that have similar learning objectives to help students gain listening and speaking skills and learning strategies to be able to communicate in everyday situations.

The results from content analysis show that the important factors that obstruct the students' language use are the errors in pronunciation and grammar, together with a limited range of vocabulary. In order to help them enhance their language proficiency, grammar, vocabulary and pronunciation needed for each language situation should be introduced in each oral class. Language fluency and complexity are the abilities that should be developed after language comprehension and accuracy have been integratively cultivated. In addition, from the results of the study, the advanced group gained the highest scores in four language components of the oral ability. In order not to discourage the poor ones in heterogeneous classes, the advanced students can help as teacher assistants when practicing the simulated dialogues or role-play activities in class. Normally, students feel at ease interacting with peers rather than with the teacher. Remedial courses or preparation courses can help prepare the students to be ready for the course before the semester begins.

In terms of the effects of the three task types in the CMLSTT on oral ability, the results of the study revealed that the scores on the answering questions task were the highest. It does not mean that the teachers have to use only this type of test to elicit students' oral performances. Actually, it is the test type for a traditional teaching method. The CMLSTT can be used as a guideline for designing a progress test depended on the language functions of any listening-speaking courses. Creating their own test tasks based on their learners' real language use, the teachers will obtained a test that truly suits their students, their background knowledge, and also stimulates the appropriate oral proficiency levels.

5.4.2 Recommendations for further studies

Since test task characteristics influence the nature of language performance (Skehan and Foster,1999), for further studies other tasks' characteristics should be compared such as difficulty levels (easy/difficult tasks), familiarity/unfamiliarity of the content and the nature of the tasks. Moreover, the researchers can also compare the differences between a structured task and a non-structured one. The test conditions such as planning time/no planning time, giving/ not giving supplementary materials are also factors that should be taken into consideration in that whether they have effects on the students' oral abilities. As Skehan (1998) suggests that task characteristics alone do not have interesting effects on oral performance, studying task characteristics under task conditions like the planning time will obtain more in-dept

results especially the characteristics of tasks, which include difficult materials and operations. Furthermore, other task types should be studied in order to find out if they have any effects on the assessment of students' oral abilities.

Apart from the learners' language abilities, their oral abilities might be affected by their personal characteristics such as their personalities, susceptibility to anxiety or speaking strategies. These variables might be included for comparison to see if personal traits like an introvert or an extrovert affect language fluency, or if there is any interaction between test-takers' personalities and task types. Their anxiety during attending the test also might be studied if it obstructs the learners' language fluency. In terms of speaking strategies, taking risks might challenge the students to try using the new sentence structures or vocabulary to create more language complexity.

Since the CMLSTT is an innovation created to assess students' four aspects of language performance, it should be compared with a paper-based test. A paper-based test might be found to be more suitable than the computer-based test in some task types, language situations or with some subjects. Furthermore, other direct tests such as an interview, face-to-face interaction or an oral presentation might be compared with the computer-based test to find the most appropriate one for each specific language context.

Another way to evaluate the efficiency of the computer-based oral testing is to interview the test-takers about their perception of this kind of testing and their suggestions for it. The students especially the ones who are not good at operating the computers, may not prefer this type of testing. Some students might want to interact with the interlocutors as their facial expressions or body language can be used as clues to guess what the interlocutors are saying. Moreover, interacting face-to-face allows the students to ask the interlocutors to clarify their questions or sentences.

REFERENCES

THAI

สุพัฒน์ สุกมลสันต์. (2547). การวิเคราะห์ข้อทดสอบแบบอัตนัยรายข้อ. **วารสารภาษาปริทัศน์** 21: 44-55.

ENGLISH

AbuSeileek, A. (2007). Cooperative vs. individual learning of oral skills in a CALL environment. **Computer Assisted Language Learning** 20(5): 493-514.

Alderson, C. (1991). Band and scores**.** In C. Alderson and B. North (eds.), **Language testing in the 1990s.** London: Macmillian Publishers.

Alderson, C. (2000). Technology in testing: The present and the future. **System** 28: 563-603.

Bachman, L. (2002). Some reflections on task-based language performance assessment. **Language Testing** 19(4): 453-476.

Bachman, L. and Palmer, A. (1996). **Language testing in practice**. Hong Kong: Oxford University Press.

Bailey, K. and Nunan, D. (2005). **Practical English language teaching: Speaking**. Singapore: McGraw-Hill Companies.

Baker, D. (1989). **Language testing: A critical survey and practical guide**. London: Edward Arnold.

Brown, H. D. (2001). **Teaching by principles: An interactive approach to language pedagogy**. New York: Addison Wesley Longman.

Brown, H.D. (2004). **Language assessment: Principles and classroom practices**. New York: Pearson Education.

Brown, J. (1996). **Testing in language programs**. New Jersey: Prentice-Hall.

Brown, J. (1997). **Computers in language testing: Present research and some future directions**. [Online]. Available from: www.llt.msu.edu/vol.1num1/brown/default.html. [2006, September 10].

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

Bygate, M., Skehan, P. ,and Swain, M. (2001). **Researching pedagogic tasks second language learning, teaching and testing**. Malaysia: Pearson Education.

Canale, M. (1983). From communicative competence to language pedagogy In J. C. Richards and R. W. Schmidt (eds.), **Language and communication**. New York: Longman Group.

Chapelle, C. (2001). **Computer applications in second language acquisition: Foundation for teaching, testing and research**. Cambridge: Cambridge University Press.

Chiu, T., Liou, H. and Yeh, Y. (2007). A study of web-based oral activities enhanced by automatic speech recognition for EFL college training. **Computer Assisted Language Learning** 20(3): 209-33.

Choi, I., Kim, K. S. and Boo, J. (2003). Comparability of a paper-based language test and a computer-based language test. **Language Testing** 20(3): 295-320.

Council of Europe. (2001). **Common European Framework of Reference for languages: Learning, teaching, assessment**. Cambridge: Cambridge University Press.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. and McNamara, T. (1999). Dictionary of language testing. In M. Milanovic (ed.). **Studies in language testing 7** Cambridge: University of Cambridge Local Examinations Syndicate and Cambridge University Press.

Douglas, D. (2000). **Assessing language for specific purposes: Theory and practice**. Cambridge: Cambridge University Press.

Douglas, D. and Hegelheimer, V. (2007). Assessing language using computer technology. **Annual Review of Applied Linguistics** 27: 115-132.

Educational testing Service. (2001). (2001-2002 ed.) **TSE and SPEAK score user guide** [Online]. Available from: http://www.toefl/tse/tseindx.html.[2007, July 23].

Educational testing Service. (2002). (2001-2002 ed.) **TSE and SPEAK score user guide.** [Online]. Available from: http://www.toefl.org/tse/tseindx.html. [2007, July 23].

Elder, C., Iwashita, N. and McNamara, T. (2002). Estimating the difficulty of oral proficiency tasks: What does the test-taker have to offer?. **Language Testing** 19(4): 347-368.

Ellis, R. (2000). Task- based research and language pedagogy. **Language Teaching Research** 4(3): 193-220.

Ellis, R. (2003). **Task-based language learning and teaching**. China: Oxford University Press.

Estaire, S. and Zanon, J. (1994). **Planning classwork: A task-based approach**. Oxford: Macmillan Heinemann.

Foster, P. and Skehan, P. (1996). The influence of planning on performance in task-based learning. **Studies in Second Language Acquisition**, 18(3): 299-324.

Foster, P., Tonkyn, A. and Wiggleworth, G. (2000). Measuring spoken language: A unit for all reasons. **Applied Linguistics**, 21(3): 354-375.

Fulcher, G. (1998). The testing of speaking in a second language. In C. Clapham and D. Corson (eds.), **Language testing and assessment: Encyclopedia of language and education.** (Vol. 7, p. 75-86). Dordrecht: Kluwer Academic Publishers.

Fulcher, G. (2000). The communicative legacy in language testing. **System** 28(4): 483-497.

Fulcher, G. (2003). **Testing second language speaking**. Malaysia: Pearson Education.

Hatch, E. and Farhady, H. (1982). **Research design and statistics for applied linguistics**. Rowley: Newbury House Publishers.

Hinkel, E. (2005). (ed.) **Handbook in second language teaching and learning**. Mahwah, N J: Lawrence Erlbaum Associates.

Hughes, R. (2002). **Teaching and researching speaking**. Great Britain: Pearson Education.

Hughes, A. (2003). **Testing for language teachers**. Cambridge: Cambridge University Press.

Hymes, D. H. (1971). Competence and performance in linguistic theory. In R. Huxley & E. Ingram (eds.), **Language acquisition: Models and methods**. London: Academic Press.

Hymes, D. H. (1972). On communicative competence. In J. B. Pride and J. Holmes (eds.), **Sociolinguistics**. Harmonsworth: Penguin Education.

Iemjinda, M. (2547). Task based learning and communicative language skills. **Seuksasart Journal, Silapakorn University** 2(1): 37-46.

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (ed.), **Planning and task performance in a second language**. Amsterdam: John Benjamins Publishing Company.

Kenyon, D. M. and Malabonga, V. (2001). Comparing the examinees' attitudes toward computer-assisted and other oral proficiency assessments. **Language Learning and Technology** 5: 60-83.

Littlewood, W. (1981). **Communicative language teaching: An introduction**. Cambridge: Cambridge University Press.

Littlewood, W. (2004). The task-based approach: Some questions and suggestions. **ELT Journal** 58(4): 319-326.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. **Language Teaching** 40: 243-249.

Long, M. H. and Crookes, G. (1993). Units of analysis in syllabus design-the case for the task. In G. Crooks and S. M.Gass (eds.), **Tasks in pedagogical contexts**. Crevedon: Multilingual Matters.

Long, M. H. and Norris, J. (2000). Task-based language teaching and assessment. In M. Byram (ed.), **Encyclopedia of language teaching**. (pp. 597-603). London: Routledge.

Luoma, S. (2004). **Assessing speaking**. Cambridge: Cambridge University Press.

Madsen, H. S. (1991). **Computer-adaptive testing of listening and reading comprehension: Research Issues and practice**. New York: Newburry House.

Mahnert, U. (1998). The effects of different lengths of time for planning on second language performance. **Studies in Second Language Acquisition** 20(1): 83-108.

McNamara, T. (2000). **Language testing**. Hong Kong: Oxford University Press.

Morrow, K. (1979). Communicative language testing: Revolution of evolution ? In C. K. Brumfit and K. Johnson (eds.), **The communicative approach to language teaching**. Oxford: Oxford University Press.

Norris , J. M., Brown, J. D., Hudson, T. D.and Bonk, W. (2002). Examinee abilities and task difficulty in task-based second language performance assessment. **Language Testing** 19(4): 395-418.

Norris , J. M., Brown, J. D., Hudson, T. D.and Yoshioka, J. (1998). Designing second language performance assessments. Honolulu: University of Hawaii Press.

Nunan, D. (2001). **Designing tasks for the communicative classroom**. Cambridge : Cambridge University Press.

Nunan , D.(2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. **TESOL Quarterly**, 37(4): 589-613.

Nunan, D. (2004). **Task-based language teaching**. Cambridge : Cambridge University Press.

Oliveira, C. P. (2004). Implementing task-based assessment in a TEFL environment. In B. L. Leaver and J. R. Willis (eds.), **Task-based instruction in foreign language education: Practice and programs**. Washington D.C.: Georgetown University Press.

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. **Language Testing** 12: 217-237.

O'Loughlin, K. (2001). The equivalence of direct and semi-direct speaking tests. In M. Milanovic and C. Weir, (eds.), **Studies in language testing 13** Cambridge: Cambridge University Press.

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

Prabhu, N. (1987). **Second language pedagogy**. Oxford: Oxford University Press.

Robinson, P. and Ross, S. (1996). The development of tasked-based assessment in English for academic purposes. **Applied Linguistics** 17(4): 455-476.

Savignon, S. (1997). **Communicative competence: Theory and classroom practice**. USA.: The McGraw-Hill Companies.

Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. **Language Testing** 11(2): 99-123.

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

Skehan, P. (2001). Task and language Performance Assessment. In M. Bygate, P. Skehan and M. Swain (eds.), **Researching pedagogic tasks second language learning, teaching and testing**. Malaysia: Pearson Education.

Skehan, P. and Foster, P. (1997). Modifying the task: The effects of surprise time and planning type on task based foreign language instruction. **Thames Valley Working Papers in English Language Teaching**, 4: 86-109.

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

Spolsky, B. (1995). **Measured words**. Oxford: Oxford University Press.

Stricker, L. J., Wider, G. Z. and Rock, D. A. (2004). Attitudes about the computer-based test of English as a Foreign language. **Computers in Human Behavior** 20: 37-54.

Tarone, E. (1985). Variability in interlanguage use: a study of style-shifting in morphology and syntax. **Language learning** 35(3): 373-405.

Teng, H.C. (2007). **The effect of task type on EFL speaking test performance**. [Online]. Available from: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019/80/28/07/86.pdf. [2007, May 30].

Underhill, N. (2003). **Testing spoken language**. Cambridge: Cambridge University Press.

Weir, C. (1993). **Understanding and developing language tests.** New York: Prentice Hall.

Weir, C. (2005). **Language testing and validation: An evidence-based approach**. New York: Palgrave Macmillian .

Weir, C. and Wu, J. (2006). Establishing test form and individual task comparability: A case study of a semi-direct speaking test. **Language Testing** 23: 167-197.

Wiggleworth, G. (2001). Influences on performance in task-based oral assessments. In M.Bygate, P. Skehan and M. Swain (eds.), **Researching pedagogic tasks second language learning, teaching and testing**. Malaysia: Pearson Education.

Williams, M. and Burden, R.L. (1997). **Psychology for language teachers: A social constructivist approach**. Cambridge: Cambridge University Press.

Willis, J. (1996). **A framework for task-based learning**. London: Longman.

APPENDICES

Appendix A : Questionnaire

Language Functions and Topics Responding to "Communicative English 1" Course

Part 1 : Personal information

Please put a tick (√) in a box according to your personal information.

☐ an English instructor
☐ an English instructor of first year students
☐ an English syllabus designer
☐ an instructor of "Communicative English 1" course

Part 2 : Listening-speaking language functions

The listening-speaking language functions below are based on Communicative English1 course objectives, the standards of English curriculum for first year university students, sets of tasks activities for listening-speaking skills by Nunan (2004) and ALTE Framework level 1.

Please select ***five*** language functions that you think are mostly needed by the first year students and select any topics that are related to those language functions or specify your own idea.

Put a tick (√) in the box of language functions and topics.

Language Functions	Topics
☐ 1. Giving personal information	☐ a. names ☐ b. age/birthday ☐ c. marital status ☐ d. hometowns ☐ e. occupations ☐ f. telephone numbers ☐ g. addresses ☐ h. families ☐ i. friends ☐ j. abilities ☐ k. interests ☐ l. leisure activities ☐ m. others……
☐ 2. Following simple instructions	☐ a. how to do things ☐ b. location of things in the house ☐ c. renting an apartment/a flat ☐ d. workplaces ☐ e. house rules ☐ f. studying (class times, assignment times) ☐ g. others……
☐ 3. Describing the locations of things	☐ a. a bedroom ☐ b. a living room ☐ c. a bathroom ☐ d. a kitchen ☐ e. a yard ☐ f. a farm ☐ g. others……

Language Functions	Topics
☐ 4. Describing people	☐ a. characteristics ☐ b. personal appearances ☐ c. personal information ☐ d. others……..
☐ 5. Describing objects	☐ a. names ☐ b. shapes ☐ c. materials ☐ d. functions ☐ e. placements/locations ☐ f. others……
☐ 6. Asking and giving simple directions	☐ a. a university ☐ b. a city hall ☐ c. a police station ☐ d. a post office ☐ e. a department store ☐ f. a food store ☐ g. a market ☐ h. a bakery ☐ i. a bookstore ☐ j. a hospital ☐ k. a bus terminal ☐ l. a train station ☐ m. others…….....
☐ 7. Ordering food and drinks	☐ a. places ☐ b. kinds of food and drinks ☐ c. preferences ☐ d. others……..
☐ 8. Describing goods in shops and asking their prices	☐ a. clothes ☐ b. stationery ☐ c. kitchen wares ☐ d. electric /household appliances ☐ e. garden tools ☐ f. others……..
☐ 9. Describing occupations	☐ a. qualifications ☐ b. responsibilities ☐ c. workplaces ☐ d. how to dress/uniforms ☐ e. tools/equipment ☐ f. others……
☐ 10. Talking about daily routine	☐ a. getting up ☐ b. having meals ☐ c. starting/ finishing studying ☐ d. starting/ finishing working ☐ e. transportation ☐ f. leisure activities ☐ g. going to bed ☐ h. others……..

Language Functions	Topics
☐ 11. Talking about past events	☐ a. what the students did ☐ b. their trips ☐ c. their vacations ☐ d. their experiences ☐ e. others……..
☐ 12. Talking about vacation plans	☐ a. destination ☐ b. length of vacation ☐ c. transportation ☐ d. accommodation ☐ e. famous places ☐ f. activities ☐ g. others……..
☐ 13. Talking on the phone	☐ a. leaving a formal/informal message ☐ b. making an appointment with a clinic/hospital ☐ c. reserving a flight ☐ d. reserving a table ☐ e. others……
☐ 14. Asking for basic services	☐ a. a bank ☐ b. a post office ☐ c. a hotel (e.g. booking a room) ☐ d. a hospital(e.g. indicating a medical problem) ☐ e. a travel agency (e.g. getting general information) ☐ f. others……
☐ 15. Expressing and sharing feelings and ideas	☐ a. sports/ events ☐ b. occupations ☐ c. leisure activities ☐ d. music/song ☐ e. movie stars/singers ☐ f. radio/ TV programs ☐ g. film/ movies ☐ h. foreign languages ☐ i. others……..

Appendix B : Specifications for the CMLSTT

The purpose of the test

The test is given to the students during the first semester of the academic year 2009 as a progress test of "Communicative English I" course which is aimed to assess the students' oral abilities in communicating in real life situations.

The test takers

The examinees are about 18-20 years old and they have learned English for at least 12 years. They are first-year undergraduate students in English and Business-English major of the English Program. They have the same language background such as number of years in studying English, no experiences in exposing to English outside class, never visiting an English-speaking country or practicing English according to their preferences.

The test constructs

The constructs to be measured include the course description and contents of "Communicative English I" course. To support the framework, the constructs of "standards of Thai University English foundation I and II" and the framework for listening and speaking skill at the beginning level (ALTE Level 1) were also explored. The constructs of the CMLSTT are given below.

Task Types	Language Functions/Constructs	Task No.
1. Answering questions	- Giving personal information	1
	- Talking about daily routine	4
2. Describing the pictures	- Giving personal information	2
	- Talking about daily routine	5
3. Transferring information	- Giving personal information	3
	- Talking about daily routine	6

Characteristics of the Setting

1. Physical characteristics :

The setting is the computer center on campus which is well - lit, comfortable and free from extraneous noise. Each test taker is provided with a PC with a speaker or a headset. He varies in the degree of familiarity with PC. Pens and paper are provided for note-taking during the preparation time.

2. Participants :

The test administrators are trained and experienced in computer-based testing. Though the relationship between the test administrators and the test takers has not been established, the test administrators have positive attitudes toward the test takers and are available to help them.

3. Time of tasks :

The test tasks are administered during the day within a fixed time period.

Characteristics of the test rubrics

1. Instructions :

The oral instructions are given in the target language through the computer. The test takers are also told about the scoring procedure and criteria. An example is provided in order to clarify what is to be done.

2. Structure :

This test is a semi-direct listening speaking test. The test takers are required to respond to the questions of three task types (answering questions, describing pictures and transferring information) in two language functions : giving personal information and talking about daily routine. The length of the answers depends on the contents the test takers have to cover and the contents are supported by the provided questions, visual aids, and guided cues. Each test item is independent from each other and it is equally weighted. The examinees have to answer the questions from the interlocutor in the video clip by recording their responses into the computer.

3. Time allotment :

The test takers are given four minutes to complete each task : one minute for watching the video clip, preparing the response for two minutes and one minute for speaking. The total time for the six tasks is about 25 minutes.

4. Scoring Method :

The criterion – referenced ability – based analytic rating scales are used to assign students' performance. The speaking abilities are judged from the accuracy, fluency, complexity and comprehension of their responses. Each item is scored based on the rating scales ranging from the scores of 1-5. The test takers' responses are rated by two experienced raters.

Characteristics of the input

1. Format

In the video clip, the input includes open-ended questions presented by the interlocutors in the target language. The activities and the language are similar to those used in their class. The test takers have to respond to the questions within the given time.

Tasks 1,4 : The input embraces the guided questions about the test takers' personal information and daily routines.

Tasks 2,5 : The input includes the pictures of the personal information and daily routines of famous people.

Tasks 3,6 : The charts of the personal information and daily routines of famous people are provided as an input in order to let the test takers talk according to the information in the charts.

2. Language

The language is general English conversation and is targeted to the beginners. The topical knowledge required in Tasks 2,3,5,6 is the personal information about famous Thai athletes and English tutors in Thailand.

Characteristics of the expected response

1. Format

The expected response consists of extended responses produced orally in the target language for the open-ended questions. In Tasks 1 and 4 the test takers do not need to have special skills in doing the test. Though the skill in describing the pictures is needed in Tasks 2 and 4 and as for the skill in transferring information from the written words, the test takers have been familiar with them since they have practiced them in class.

2. Language

Tasks 1-3 : The vocabulary included in the expected response is concerned with personal information such as names, hometowns, families, occupations or interests. The vocabulary is expressed in the sentence structures that are familiar to the students.

Tasks 4-6 : The expected response in this task contains general vocabulary necessary for describing the daily routines of the people. The sentence structures are frequently used in class.

Relationship between input and response

The test takers are expected to process non-reciprocal responses, i.e. their answers respond to the given input but do not create other interactions after this. The responses are based on both direct and indirect relationship with the input. The direct relationship between the input and response occurs in Tasks 1 and 4. Though there are some guided questions given to the test takers, they have to produce their own answers according to their personal information and daily routines. The test takers must rely on their own knowledge of how to describe personal information and daily routines that are familiar to them. While in Tasks 2,3,5 and 6 the test takers have to use the given information from the pictures or information in the chart to produce their sentences and use their own topical knowledge to add some more information about the people.

Appendix C : Task Congruence Rating Result (CMLSTT Validation Forms)

Task No.	Contents	Judges				
		1	2	3	4	5
1	1. Objective : To give personal information by answering the provided questions	0	1	1	1	1
	2. Topic : Myself	1	1	1	1	-
	3. Function : Giving personal information	1	1	1	1	-
	4. Question Type : - Polite request - Wh – question	1	1	0	0	-
	5. Difficulty Level : beginner	1	1	1	1	-
2	1. Objective : To give personal information by using the provided pictures and cues	0	1	1	1	1
	2. Topic : Myself	-1	1	0	0	-
	3. Function : Giving personal information	1	1	1	1	-
	4. Question Type : polite request	1	1	0	1	-
	5. Difficulty Level : beginner	1	1	1	1	-
3	1. Objective : To give personal information by using the provided words in the chart	0	1	1	1	1
	2. Topic : Myself	-1	1	0	0	-
	3. Function : Giving personal information	1	1	1	1	-

Task No.	Contents	Judges				
		1	2	3	4	5
3	4. Question Type : polite request	1	1	0	1	-
	5. Difficulty Level : beginner	1	1	1	1	-
4	1. Objective : To tell the daily routine by answering the provided questions	0	1	1	1	1
	2. Topic : My life schedule	0	1	1	1	-
	3. Function : Telling the daily routine	1	1	1	1	-
	4. Question Type : Wh-questions	1	1	0	1	-
	5. Difficulty Level : beginner	1	1	1	1	-
5	1. Objective : To tell the daily routine by describing the provided pictures using cued words	0	1	1	1	1
	2. Topic : My life schedule	-1	1	1	0	-
	3. Function : Telling the daily routine	1	1	1	1	-
	4. Question Type : Wh-questions	1	1	0	1	-
	5. Difficulty Level : beginner	1	1	1	1	-
6	1. Objective : To tell the daily routine by using the provided words in the chart	0	1	1	1	1
	2. Topic : My life schedule	-1	1	1	1	-
	3. Function : Telling the daily routine	1	1	1	1	-
	4. Question Type : Wh-questions	0	1	0	1	-
	5. Difficulty Level : beginner	1	1	1	1	-

Appendix D : The average scores of the six tasks

The average scores arranged from those of the highest to the lowest students in doing each task and used to calculate the discrimination and difficulty indexes

	task 1	task 2	task 3	task 4	task 5	task 6
student 1	**19**	**19.5**	**19**	**18.5**	**18**	**18**
student 2	**18.5**	**18**	**18.5**	**18**	**17.5**	**16.5**
student 3	**16**	**16.5**	**16.5**	**16.5**	**16**	**16**
student 4	**15**	**16**	**16**	**16.5**	**15.5**	**16**
student 5	**15**	**15**	**15.5**	**16.5**	**15.5**	**15.5**
student 6	**14.5**	**14**	**14.5**	**16**	**15**	**14.5**
S_H	**98**	**99**	**100**	**102**	**97.5**	**96.5**
student 7	14.5	13.5	13	15.5	14	13.5
student 8	14	13	13	15.5	12.5	13
student 9	13.5	12.5	12.5	14.5	12	12.5
student 10	12.5	12	12.5	14.5	11.5	12
student 11	12.5	11.5	12.5	14.5	11.5	12
student 12	12	11	12	13	11.5	11.5
student 13	12	11	11.5	12.5	11.5	10.5
student 14	11.5	11	11.5	12.5	11.5	10.5
student 15	11	10.5	10.5	12.5	11.5	10.5
student 16	11	10.5	10.5	12.5	10.5	10.5
student 17	11	9.5	10.5	12	10.5	10
student 18	10.5	9.5	10.5	12	9.5	9
student 19	**9.5**	**9.5**	**10**	**11.5**	**9.5**	**9**
student 20	**9.5**	**9.5**	**9**	**10.5**	**9**	**8.5**
student 21	**9**	**8.5**	**8**	**10**	**7.5**	**8**
student 22	**9**	**5.5**	**7**	**9.5**	**7**	**7**
student 23	**8**	**5.5**	**7**	**9.5**	**6**	**7**
student 24	**6**	**5**	**6**	**6.5**	**5.5**	**5**
S_L	**51**	**43.5**	**47**	**57.5**	**44.5**	**44.5**

Appendix E : The mean scores and standard deviations of six tasks

Frequencies

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
N	24	24	24	24	24	24
Mean	12.29	11.58	11.98	13.38	11.67	11.52
S.D.	3.165	3.726	3.447	2.979	3.416	3.383

Appendix F : Reliability of CMLSTT

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.972	.973	6

Case Processing Summary

		N	%
Cases	Valid	24	100.0
	Excluded[a]	0	.0
	Total	24	100.0

a. Listwise deletion based on all variables in the procedure.

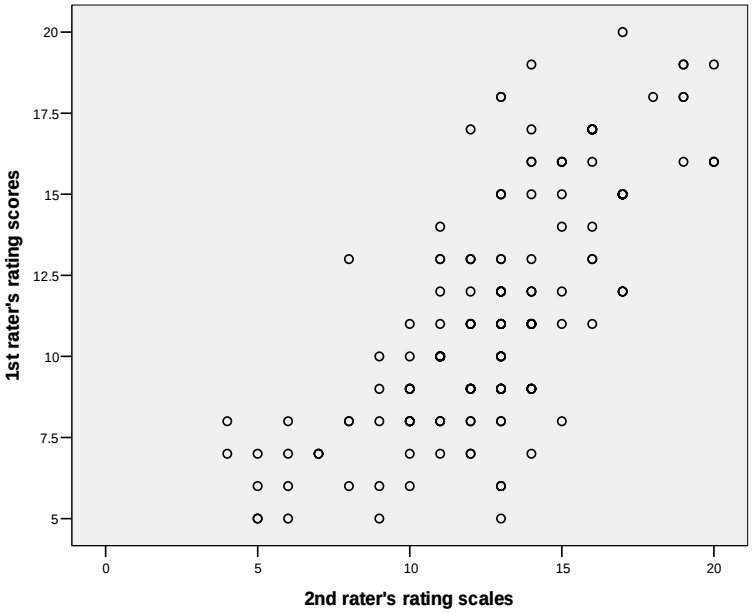
Inter-Item Correlation Matrix

	task1 scores	task2 scores	task3 scores	task4 scores	task5 scores	task6 scores
task1 scores	1.000	.915	.819	.886	.833	.818
task2 scores	.915	1.000	.876	.892	.904	.846
task3 scores	.819	.876	1.000	.810	.932	.838
task4 scores	.886	.892	.810	1.000	.848	.831
task5 scores	.833	.904	.932	.848	1.000	.829
task6 scores	.818	.846	.838	.831	.829	1.000

The covariance matrix is calculated and used in the analysis.

Appendix G(1) : Scatter Plot and the Correlation Coefficient

Scatter Plot Representing the Scores by Two Raters and the Correlation Coefficient between Two Raters's rating (1st round)



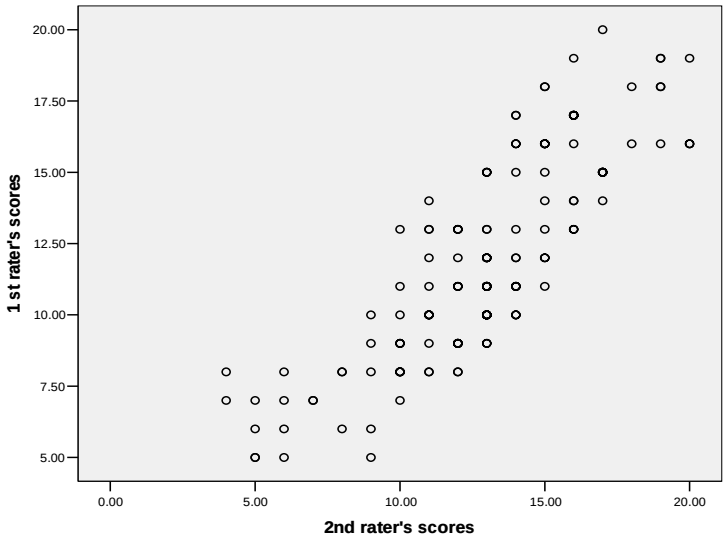
Correlations

		1st rater	2nd rater
1st rater's scores	Pearson Correlation	1	.735**
	Sig. (2-tailed)		.000
	N	144	144
2nd rater's scores	Pearson Correlation	.735**	1
	Sig. (2-tailed)	.000	
	N	144	144

** Correlation is significant at the 0.01 level (2-tailed).

Appendix G(2) : Scatter Plot and the Correlation Coefficient

Scatter Plot Representing the Scores by Two Raters and the Correlation Coefficient between Two Raters' Rating (2nd round)



Correlations

		1st rater	2nd rater
1st rater's scores	Pearson Correlation	1	.833**
	Sig. (2-tailed)		.000
	N	144	144
2nd rater's scores	Pearson Correlation	.833**	1
	Sig. (2-tailed)	.000	
	N	144	144

** Correlation is significant at the 0.01 level (2-tailed)

Appendix H : Computer-mediated Listening-Speaking Test Tasks

Directions

This listening-speaking test is composed of **3 tasks** based on the following **3 task types** :

1. **answering questions** (Tasks **1**)
2. **describing pictures** (Tasks **2**)
3. **transferring information** (Tasks **3**)

Each task type contains **the language function** (giving personal information). In each listening–speaking task, you will see a video clip of the situation in which the conversation takes place. You have to assume yourself as a speaker in that situation. The interlocutor in the video will ask you a short question. You can take notes while listening and use these notes together with the provided questions, visual aids and charts to answer each question.

This test will last **approximately 15 minutes**. You are given four minutes to complete each task: **1 minute** to listen to the direction and watch the video clip, **2 minutes** to prepare your response and **1 minute** to speak.

Answer each question by speaking into a microphone and record your voice into the computer. Your score will be based on your recorded speech, so be sure to **speak loudly** enough to record what you say. Your spoken response will be sent to 2 experienced raters to score your speaking ability.

If you have any questions, please ask the test administrator now.

Thank you for your cooperation.

Task 1

Narrator

Watch the video clip and introduce yourself by answering all of the provided questions. You will hear each question twice. After you hear all questions, you will have 2 minutes to prepare your responses and 1 minute to speak.

Man : Good morning, I'm David Baxie. I would like to get some information about you. Please answer the following questions and give additional details to your answers.

1. When were you born?
2. How many brothers and sisters do you have?
3. What do your father and mother do?
4. Where is your hometown?
5. What do you like to do in your free time?

Narrator

You now have 2 minutes to prepare your response. Start speaking after you hear the beeping sound.

Task 2

Narrator

Suppose your school invites Paradon to be its guest. As one of the teachers, you have to introduce him to your students. Watch the video clip about the biography of "Paradon Srichapan" and answer the question by describing his personal information. Give details for each picture. You will have 2 minutes to prepare your response and 1 minute to speak.

1st teacher : I've heard that we have a special guest to our school. Who is he?

Narrator

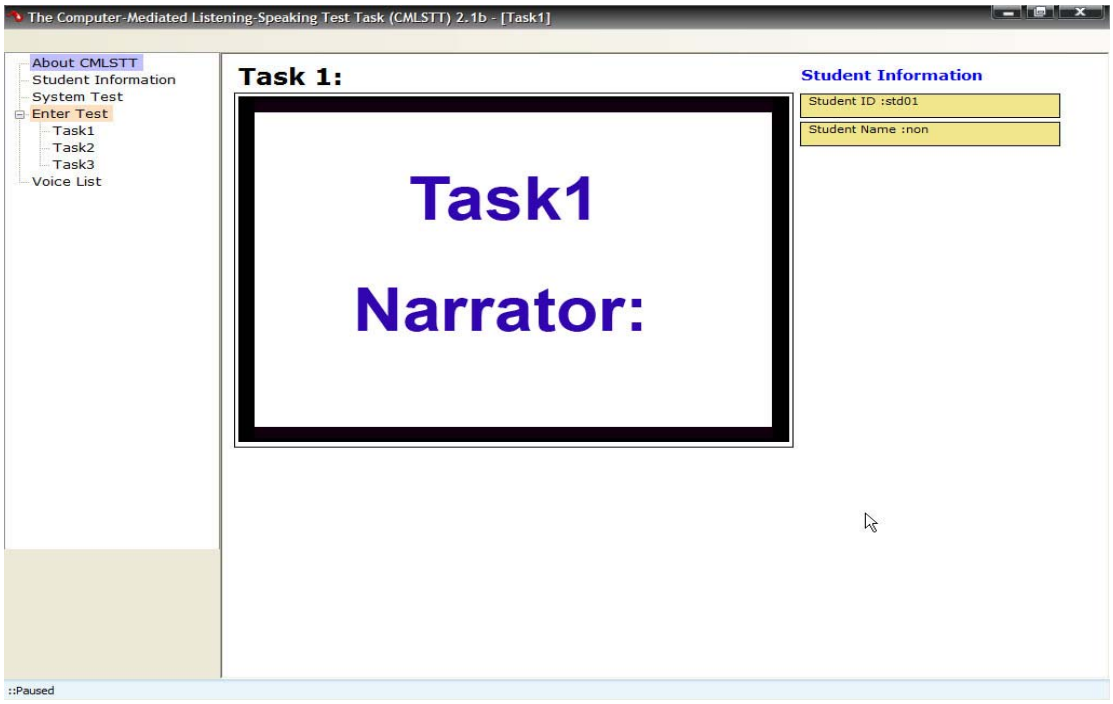
You now have 2 minutes to prepare your response. Start speaking after you hear the beeping sound.

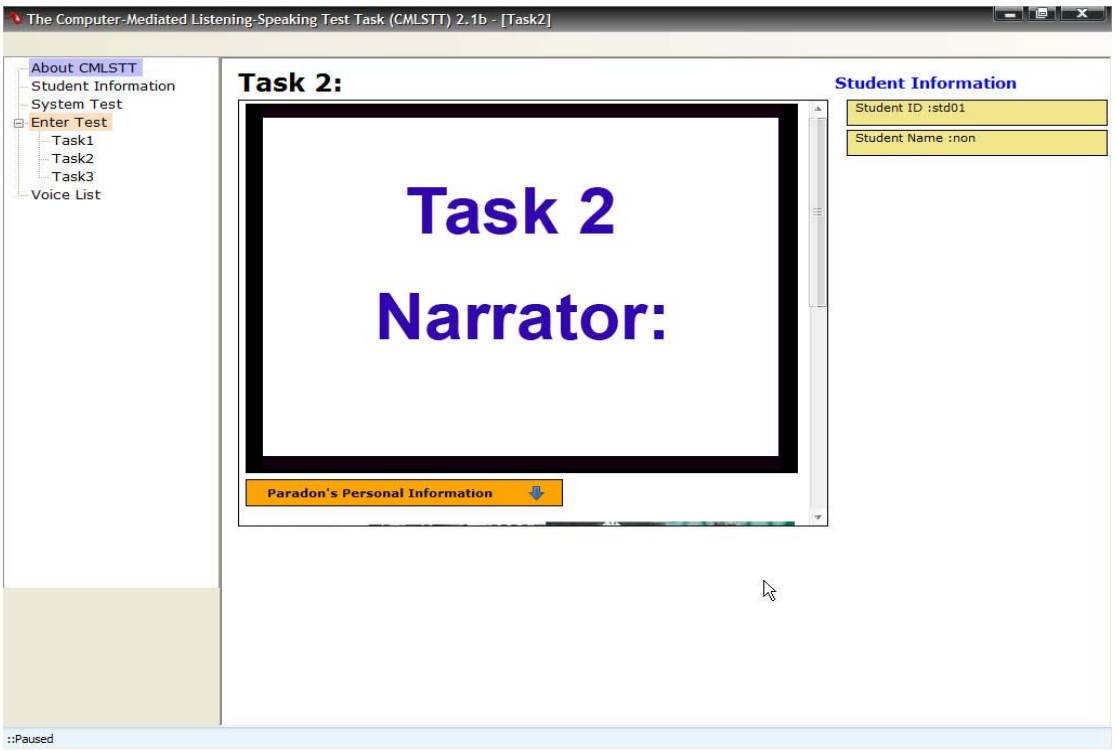
Task 3

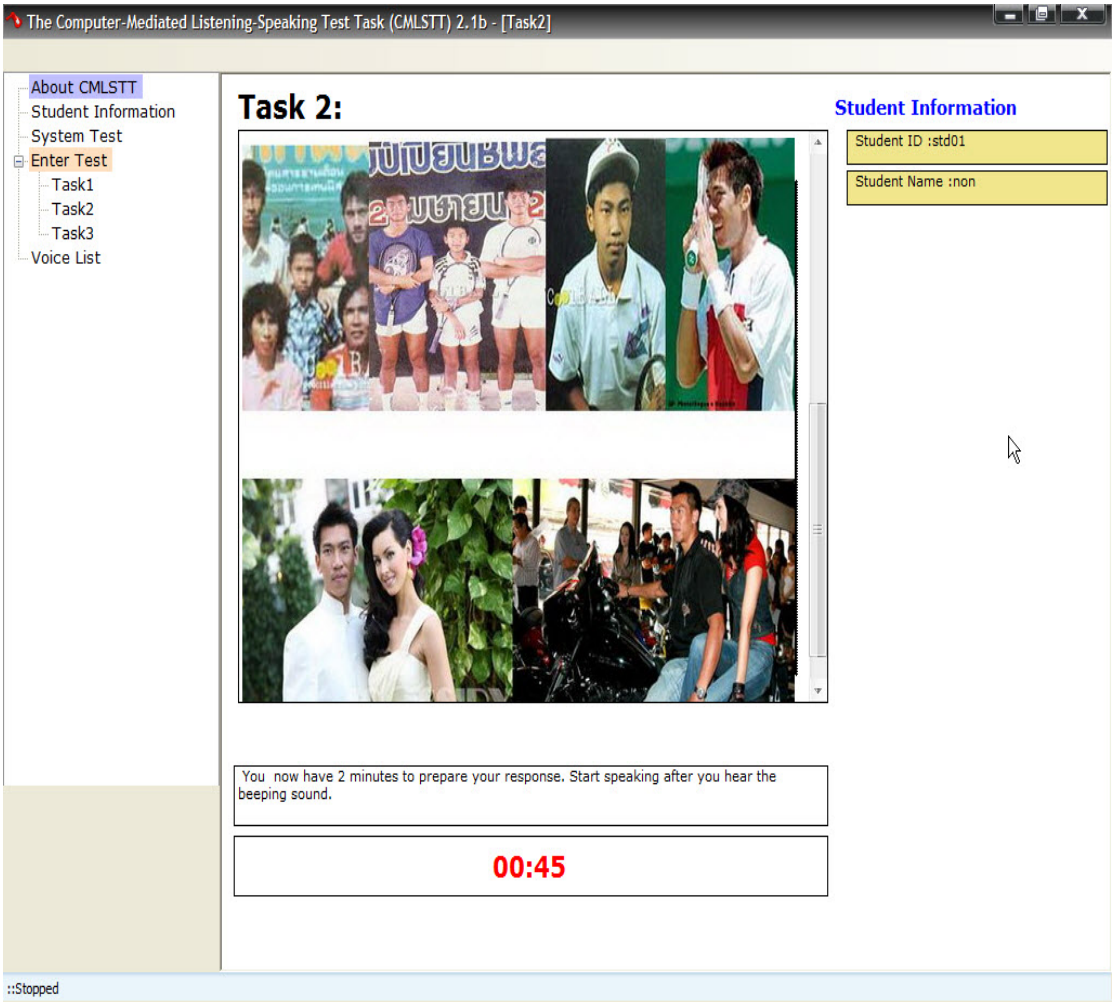
Narrator

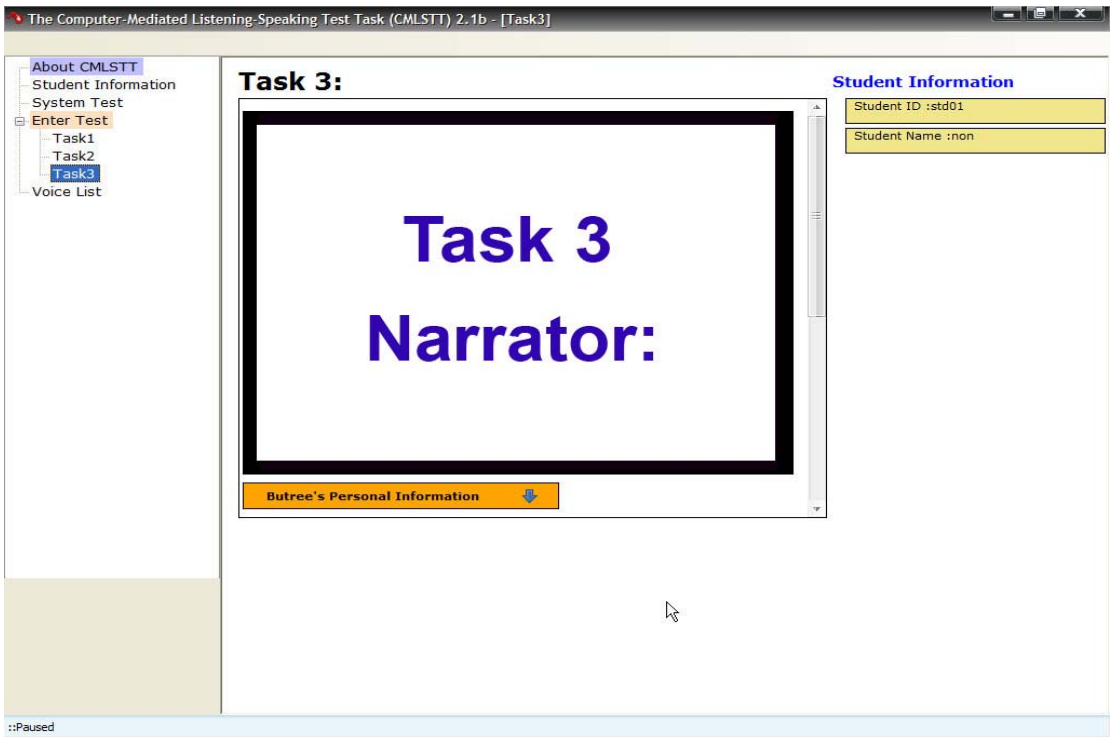
Suppose your foreign friend asks you about an Olympic 2008 Silver Medalist in Taekwando, Butree Puedpong. Answer the question by describing her personal information based on the cued words in the form. You will have 2 minutes to prepare your response and 1 minute to speak.

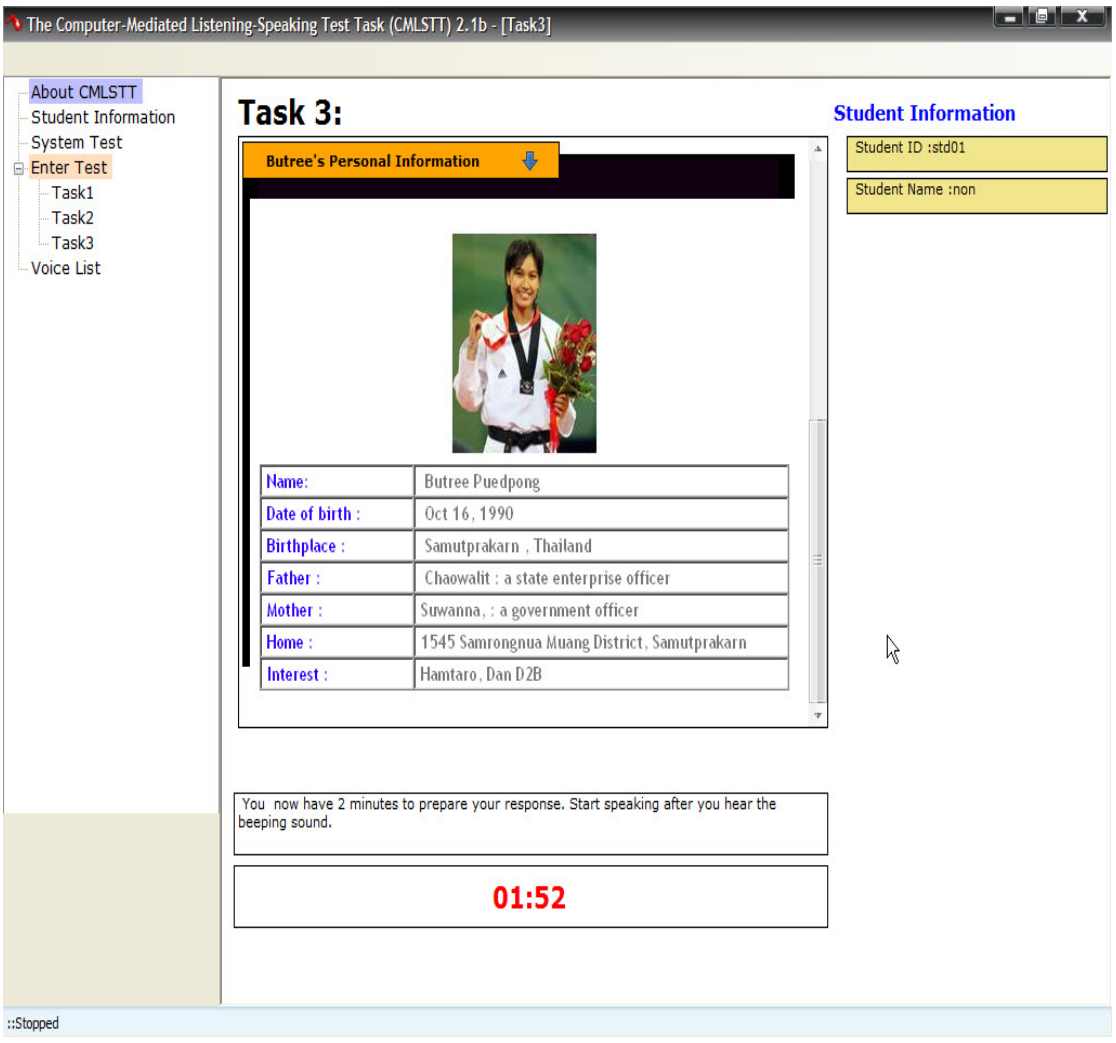
Your friend : Who is the girl winning the silver medal for Taekwando?


The Computer-Mediated Listening-Speaking Test Task (CMLSTT) 2.1b - [Task1]
About CMLSTT
Student Information
System Test
Enter Test
Task1
Task2
Task3
Voice List
Task 1:
Task1
Narrator:
Student Information
Student ID :std01
Student Name :non
::Paused


The Computer-Mediated Listening-Speaking Test Task (CMLSTT) 2.1b - [Task2]
About CMLSTT
Student Information
System Test
Enter Test
Task1
Task2
Task3
Voice List
Task 2:
Task 2
Narrator:
Paradon's Personal Information
Student Information
Student ID :std01
Student Name :non
::Paused


The Computer-Mediated Listening-Speaking Test Task (CMLSTT) 2.1b - [Task2]
About CMLSTT
Student Information
System Test
Enter Test
Task1
Task2
Task3
Voice List
Task 2:
Student Information
Student ID :std01
Student Name :non
You now have 2 minutes to prepare your response. Start speaking after you hear the beeping sound.
00:45
::Stopped


The Computer-Mediated Listening-Speaking Test Task (CMLSTT) 2.1b - [Task3]
About CMLSTT
Student Information
System Test
Enter Test
Task1
Task2
Task3
Voice List
Task 3:
Student Information
Student ID :std01
Student Name :non
Task 3
Narrator:
Butree's Personal Information
::Paused


The Computer-Mediated Listening-Speaking Test Task (CMLSTT) 2.1b - [Task3]
About CMLSTT
Student Information
System Test
Enter Test
Task1
Task2
Task3
Voice List
Task 3:
Student Information
Student ID :std01
Student Name :non
Butree's Personal Information
Name: Butree Puedpong
Date of birth : Oct 16, 1990
Birthplace : Samutprakarn , Thailand
Father : Chaowalit : a state enterprise officer
Mother : Suwanna, : a government officer
Home : 1545 Samrongnua Muang District, Samutprakarn
Interest : Hamtaro, Dan D2B
You now have 2 minutes to prepare your response. Start speaking after you hear the beeping sound.
01:52
::Stopped

Appendix I : The transcript of the three groups of students in doing 3 tasks.

The transcript of *the advanced, intermediate and beginning groups* in doing *task 1* (answering questions)

The advanced group

Std 19

Hello. My name is Siriporn Moonkao. You can call me Gif. I'm nineteen years old. I'm studying in the first year of Business English Program. For the answer

The first one : I was born in twenty of August, nineteen ninety.

The second one : I have two sister and one brother.

The third one : My mother is a housewife. My father is a farmer.

The fourth one : My hometown is in Pakonchai.

The last one : In my free time, I like watching TV. and play the internet.

Std 21

My name is Wasan Karunram. My birthday is June eighteen, nineteen ninety. I live in Buriram Province, Thailand. I have one younger sister. My father is Somsak, he is a farmer. My mother is Lampueng, he is a farmer,too. My address is one-eight-two Chumhet, Muang District, Buriram Province, Thailand. My free time, I like listen to music.

Std 02

I was born on the fifth of December, nineteen eighty-nine. I have only one brother. He is six years older than me and my mother is a house wife and my father is a teacher and my hometown is Buriram. On my free time, I like to watch TV. and I like to ply sport play badminton.

Std 16

1. I was born on twenty-seven of November, nineteen ninety.
2. I no have brother and sister. I am an only child.
3. My father is business, personal business but my mother is a housewife.
4. My hometown is Tumbon Banyang, Amphur Muang, Changwat Buriram.
5. In my free time, I listen to the music and use the internet.

Std 18

I born in the third December in nineteen eighty-six. I no have any brother and sister. My father and mother have do a farmer. My hometown in Surin. On my free time I like listen to music and go shopping with my friend.

The intermediate group

Std 44

My name is Sirirat Patiram. My birthday twenty-one March, nineteen eighty-eight. I have sister two and none brother. Father and mother are farmer. I live in Buriram. I read book and listen and play a game.

Std 42

My name is Anongnat Saothong. I am from Buriram. I have one brother and me. My parents is …?..... .My free time I like reading listen to music. My family is small. There are four people in family.

Std 49

My name is Somporn Detkulram. My nickname Oh. Birthday three May oe nine hundred and ninety. I have one brother and one sister. My mother name is Sawai Detkulram. My father name is Saming Detkulram. I live in town Koomuang and I live in Buriram.

Std 46

I have one sister and one brother. My father, he is die but my mother is teacher. My home.. my house.. my house I live in Prakonchai. My free time I play basketball everyday.

Std 37

I born thirty May, one thousand ninety. Two people my father and my mother they are farmers. I live in Nong Ki. I read a book.

The beginning group

Std 73

S-R-I-S-U-D-A N-O-N-K-O-N-G My sister is one sister. Mother and father we are farmers.

Std 64

- Twenty-seven May, nineteen ninety
- I have young one brother.
- Buriram
- Play basketball.

Std 76

I have one brother mother father. I live fifty-nine …..?..... Amphur Muang, Buriram.

Std 62

I was born twelve December and nine ten year old.

Std 67

One brother and one brother

The transcript of *the advanced, intermediate and beginning groups* in doing *task* 2 (Describing Pictures)

The advanced group

Std 16

He play tennis since he was a child. Then he is the first of Thailand and the first of the world. He's married with Natalie and thus he sit on the motorcycle with Natalie.

Std 20

Paradon is playing tennis. He is the one of Thai. She play tennis every day in tennis player. He marry with Natalie Glebova. She is Miss Universe. He is very happy and hards working. He is tall and handsome. Natalie is beautiful and Paradon is handsome.

Std 18

Paradon have to play tennis when he young and still play tennis now. And after that Paradon is marries with Natalie. She name Natalie from Canada. (pause) Paradon and Natalie sit on motorbike together.

Std 27

He is Paradon. Paradon he have one brother. He like play tennis because my father of Paradon play tennis. He married Natalie because Paradon love Natalie. He start playing tennis begin children. He like to play tennis and he love Natalie very much. He not work and he play tennis.

Std 19

His name is Paradon Srichapan. His hometown is in Kon Kaen. He play tennis since he was young. He is tennis player. He is married with Miss Universe. Her name is Natalie Glebova.

The intermediate group

Std 35

Paradon Srichapan is champion of tennis of Thailand. And she married with Natalie Glebova. (Pause for quite a while) is Miss Universe.

Std 37

In picture has Paradon Srichapan. He is a …?..... .play he play tennis. He has family. In the family have five people. He love Natalie Glebova. He like tennis…. very much.

Std 48

Paradon is a tennis. Your like tennis. Your have girlfriends name is Natalie. Paradon with the winner champion the tennis. Paradons marry…married Natalie in Thailand.

Std 32

OK. Paradon Srichapan he is a ……..?..... hes play table.. uh..tennis sport. He is a master ..uh..sport..uh.. tennis sport in Thailand. He's a he's marry uh he's marry Natalie. He's champion…he is champion tennis of the world.

Std 58

He is Paradon. He is play tennis. He married is Natalie. (Pause) And he play tennis in the Thailand.

The beginning group

Std 84

Paradon is a badminton sport. He live in Thailand. My family is a Paradon have a happy. Paradon stay in at Natalie. She is a happy to a family. My family is a sport. He is a badminton sport.

Std 86

Paradon is play tennis Okexercise(Thai word).

Std 74

Paradon is very excellent.

Std 76

Paradon is married to Natalie. He play tennis or he ……?.....tennis.

Std 79

Paradon like tennis sport.

The transcript of *the advanced intermediate and beginning groups* in doing *task* 3 (Transferring information)

The advanced Group

Std 21

Her name is Butree. She was born on sixteen of October, nineteen ninety. She live in Samutprakarn, Thailand. Her father name is Chaowalit, he is state enterprise officer. Her mother's name is Suwanna. She is a government officer. Her address is one five four five Samrong Nua, Samutprakarn. She like Hamtaro and Dan D2B.

Std 19

Her name is Butree Puedpong. She was born on sixteen of October, nineteen ninety. She was born at Samutprakarn. Her father his name is Chaowalit. He work in the office. Her mother name is Suwanna. She work at the government officer. Now she live in one five four five at Samrong Nua, Muang Samutprakarn. Her interest are Hamtaro and Dan D2B.

Std 02

The woman in this picture, her name is Butree. She was born on sixteen of October, nineteen ninety. She live in Samutprakarn, Thailand. Her father name is Chaowalit, he is state enterprise officer. Her mother's name is Suwanna. She is a government officer. Her address is one five four five Samrong Nua, Samutprakarn. She like Hamtaro and Dan D2B.

Std 01

In the picture she name is Butree. Her birthday is sixteen October, nineteen ninety. Her birthplace is Samutprakarn, Thailand. Her father's name is Chaowalit, he is a state enterprise officer. She name is Suwanna. He is a government officer.

Std 20

Miss Butree Puedpong . She birth of sixteen of October in nineteen ninety. She birthplace at Samutprakarn in Thailand. Father is name Chaowalit, she is a state enterprise officer and mother name Suwanna. She is a government officer. She live at one five four five Samrong Nua, Muang District, Samutprakarn. She like Hamtaro and idol she like Dan D2B.

The intermediate Group

Std 37

She name is Butree Puedpong. Birthday October sixteen one thousand ninety. She live in Samutprakarn in Thailand. My father she name Chaowalit. My mother she name Suwanna. She live in one thousand five forty-five, Samrong Nua, Muang District, Samutprakarn.

Std 38

Butree personal information. Butree Puedpong date of birth sixteen nineteen ninety. Birthplace Samutprakarn , Thailand. Farmer is Chaowalit, a state enterprise officer. Mother is Suwanna, a government officer. Interest Hamtaro, Dan D2B. Butree play taekwando.

Std 47

Butree date of birth sixteen October nineteen ninety. Birthplace, Samutprakarn in Thailand**.** My father she name is Chaowalit a state enterprise officer. My mother she name is Suwanna a government officer. Home sixteen forty-five Samrongnua, Muang District, Samutprakarn. Interest Hamtaro and Dan D2B.

Std 48

Her name Butree. Date of birth October sixteen. Birthplace Samutprakarn , Thailand**.** Butree father is Chaowalit. Mother name is Suwanna. Home interest Dan D2B.

Std 44

Butree Puedpong date of birth October sixteen, nineteen ninety. Birthplace Samutprakarn , Thailand**.** Father Chaowalit a state enterprise officer. Mother Suwanna a government officer. Home fifteen fourteen Samrongnua District, Samutprakarn. Interest Hamtaro and Dan D2B. She is beautiful and very hot.

The beginning Group

Std 66

Butree Puedpong…date of birth October sixteen nineteen ninety…birthplace, I don't know…father, I don't know…mother Suwanna…home one five four five… interest Hamtaro, Dan D2B

Std 87

You have sister. You live in Thailand. You beautiful and lovely. You like sport in Thailand.

Std 79

Butree like taekwando sport. She was born in October nineteen ninety. She live in Thailand.

Std 74

She name is Butree Puedpong. Date of birth October sixteen nineteen ninety.

Std 73

She name is Butree Puedpong. She live in Samutprakarn in Thailand.

Appendix J

CMLSTT Rating Scales

Scoring is based on four aspects of language performance : accuracy ,fluency, complexity, and comprehension. The subskills of each aspect are judged from the students' abilities and skills in performing the tasks.

Level	Accuracy	Fluency	Complexity	Comprehension
4.1-5.0	show a high degree of grammatical accuracy with adequate vocabulary to convey the meaning precisely to complete the task ; use appropriate stress and intonation patterns and rarely make any mistakes	be able to speak intelligibly and produce smoothly flowing, natural and coherent speech	give very detailed responses by using complex grammatical structures with a wide range of connectors and cohesive devices	can understand the instruction and situation well by giving adequate and relevant responses to complete the task
3.1-4.0	maintain a high degree of grammatical control and accurate vocabulary; use appropriate stress and intonation patterns but sometimes make mistakes that do not cause misunderstanding or can be corrected when occur	talk fluently and spontaneously ; in spite of a few hesitations and pauses, the students can maintain a flow of speech	give clear responses by using well-structured speech with a range of connectors and cohesive devices	can understand most of the instruction and situation by give relevant responses needed to complete the task

CMLSTT Rating Scales

Level	Accuracy	Fluency	Complexity	Comprehension
2.1-3.0	be able to speak the language with sufficient vocabulary in simple sentence structures; accent is intelligible but sometimes make basic mistakes in vocabulary and grammar that interfere understanding	produce utterances rather slowly but smoothly; some hesitations occur while the students find the suitable words and structures, a few long pauses can be noticed but they can keep on talking	use simple and complex structures with simple connective markers repetitively	can understand the instruction and situation quite well but sometimes give inadequate and irrelevant responses to complete the task
1.1-2.0	show only limited control of vocabulary and a few simple grammatical structures; errors in pronunciation, inaccurate vocabulary and grammatical mistakes interfere understanding	speech rate is rather slow and contains frequent long pauses and hesitations while searching for less familiar words; the sentences are disjointed and limited in length	use a limited range of vocab in incomplete structures and can link groups of words with simple connectors	can understand some instruction and situation and then give generally inadequate and irrelevant responses to complete the task

179

CMLSTT Rating Scales

Level	Accuracy	Fluency	Complexity	Comprehension
0.1-1.0	use short and incomplete sentences with isolated words that can express intentions at only elementary needs because of inaccurate vocabulary and grammatical patterns; errors in pronunciation, vocabulary and grammar are frequent	use a few short sentences and no fluent speech can occur	very low-level responses rely on isolated words or short utterances and can link words or groups of words with very basic connectors	cannot understand the instruction and situation then cannot perform the function

180

BIOGRAPHY

Saowarot Ruangpaisan was born in 1964 in Bangkok. She received a Bachelor of Education, majoring in English from Srinakarinwirot University (Prasarnmit Campus) in 1986. She obtained her Master of Education in Teaching English as a Foreign Language (TEFL) from Silapakorn University (Sanamchan Campus), Nakornpathom in 1994. She has been a lecturer in the department of English at Buriram Rajabhat University for twelve years. Her interests lie in the areas of English teaching methodology, teaching grammar and language testing.