

THE EFFECTS OF WEB-BASED SPEAKING TEST IN ENGLISH  
FOR TOURISM (WBST-EFT) TASK TYPES AND ENGLISH  
PROFICIENCY LEVELS ON STUDENTS' SPEAKING  
PERFORMANCE AND AN INVESTIGATION OF THEIR  
ATTITUDES AND TEST TAKING STRATEGIES

Ms. Malinee Phaiboonnugulkij

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 2011

Copyright of Chulalongkorn University  
บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่เก็บถาวรในคลังปัญญาจุฬาฯ (CUIR)

เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)  
are the thesis authors' files submitted through the Graduate School.

ผลกระทบของชนิดกิจกรรมในแบบทดสอบการพูดบนเว็บไซต์สำหรับภาษาอังกฤษเพื่อการ  
ท่องเที่ยวและระดับความสามารถภาษาอังกฤษที่มีต่อความสามารถของนักศึกษาในการทำ  
แบบทดสอบการพูดและการศึกษาเจตคติและกลวิธีในการทำข้อสอบการพูด

นางสาวมาลินี ไพบูลย์นุกุลกิจ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต

สาขาวิชาภาษาอังกฤษเป็นภาษานานาชาติ (สหสาขาวิชา)

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2554

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title THE EFFECTS OF WEB-BASED SPEAKING  
TEST IN ENGLISH FOR TOURISM (WBST-EFT)  
TASK TYPES AND ENGLISH PROFICIENCY  
LEVELS ON STUDENTS' SPEAKING  
PERFORMANCE AND AN INVESTIGATION OF  
THEIR ATTITUDES AND TEST TAKING  
STRATEGIES

By Miss Malinee Phaiboonnugulkij  
Field of Study English as an International Language  
Thesis Advisor Professor Emeritus 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 Emeritus Kanchana Prapphal, Ph.D.)

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

.....Examiner  
(Ajarn Sutthirak Sapsirin, Ph.D.)

.....External Examiner  
(Assistant Professor Somsak Boonsathorn, Ph.D.)

มาลินี ไพบูลย์นุกุลกิจ:ผลกระทบของชนิดกิจกรรมในแบบทดสอบการพูดบนเว็บไซต์สำหรับภาษาอังกฤษเพื่อการท่องเที่ยวและระดับความสามารถภาษาอังกฤษที่มีต่อความสามารถของนักศึกษาในการทำแบบทดสอบการพูด และการศึกษาเจตคติและกลวิธีในการทำข้อสอบการพูด. (THE EFFECTS OF WEB-BASED SPEAKING TEST IN ENGLISH FOR TOURISM (WBST-EFT) TASK TYPES AND ENGLISH PROFICIENCY LEVELS ON STUDENTS' SPEAKING PERFORMANCE AND AN INVESTIGATION OF THEIR ATTITUDES AND TEST TAKING STRATEGIES)

อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ศ.กิตติคุณ ดร. กาญจนา ปราบพาล, 285 หน้า.

การวิจัยนี้มีวัตถุประสงค์เพื่อ (1) ศึกษาผลกระทบของประเภทกิจกรรมในแบบทดสอบการพูดบนเว็บไซต์สำหรับภาษาอังกฤษเพื่อการท่องเที่ยวและระดับความสามารถภาษาอังกฤษที่มีต่อความสามารถของนักศึกษาในการทำแบบทดสอบการพูด ในด้านความรู้ทางภาษา ความรู้เชิงเนื้อหา และความคล่องแคล่วในการใช้ภาษา (2) ศึกษาเจตคติของนักศึกษาต่อแบบทดสอบการพูด (3) ศึกษาและเปรียบเทียบประเภท และความถี่การใช้กลวิธีในการทำข้อสอบการพูดของนักศึกษาที่มีความสามารถภาษาอังกฤษระดับสูง และต่ำ กลุ่มตัวอย่างในการวิจัยนี้เป็นนักศึกษา ชั้นปีที่ 3 ที่ลงทะเบียนเรียนรายวิชาภาษาอังกฤษเพื่อการท่องเที่ยว 2 มหาวิทยาลัยราชภัฏนครราชสีมา จำนวน 120 คน ที่ได้จากการสุ่มแบบช่วงชั้น โดยแบ่งออกเป็นกลุ่มที่มีระดับความสามารถทางภาษาอังกฤษสูง และต่ำ เครื่องมือที่ใช้ในการวิจัยประกอบด้วย แบบสอบถามความต้องการในบริบทของการท่องเที่ยว แบบทดสอบการพูดบนเว็บไซต์สำหรับภาษาอังกฤษเพื่อการท่องเที่ยว แบบสอบถามเจตคติออนไลน์ที่มีต่อแบบทดสอบการพูด และแบบจำแนกประเภทกลวิธีในการทำข้อสอบการพูด ผู้วิจัยใช้การวิเคราะห์ความแปรปรวนแบบสองทาง (Two-way ANOVA) สำหรับการวิเคราะห์ผลกระทบหลักและร่วม ของระดับความสามารถทางภาษาอังกฤษและประเภทแบบทดสอบที่มีต่อความสามารถในการทำข้อสอบการพูด นอกจากนี้ใช้ t-test แบบอิสระ (Independent samples t-test) ในการวิเคราะห์และเปรียบเทียบเจตคติระหว่างนักศึกษาที่มีความสามารถทางภาษาสูง และ ต่ำต่อแบบทดสอบการพูด ใช้สถิติพรรณนา การแจกแจงความถี่ และเปอร์เซ็นต์เพื่ออธิบายและเปรียบเทียบ ส่วนประกอบการพูดภาษาเฉพาะทาง เจตคติของนักศึกษา และกลวิธีในการทำข้อสอบ รวมถึงการวิเคราะห์เนื้อหาจากผลทดสอบการพูดและจากการสัมภาษณ์แบบปากเปล่าเกี่ยวกับกลวิธีในการทำข้อสอบการพูดเพื่อให้ได้ข้อมูลเชิงลึก

ผลการศึกษาพบว่า (1) ระดับความสามารถทางภาษาอังกฤษมีผลกระทบต่อความสามารถทางการพูดภาษาเฉพาะทางอย่างมีนัยสำคัญทางสถิติ นอกจากนี้พบว่าประเภทแบบทดสอบมีผลกระทบต่อการใช้คำศัพท์ ประเภทภาษา และความรู้เชิงเนื้อหา โดยเฉพาะอย่างยิ่งประเภทแบบทดสอบมีผลกระทบต่อนักศึกษาที่มีความสามารถภาษาอังกฤษระดับต่ำมากกว่ากลุ่มที่มีความสามารถสูงในด้านความรู้เชิงเนื้อหา จาก การวิเคราะห์เนื้อหาพบลักษณะเฉพาะของส่วนประกอบการพูดที่เกี่ยวข้องกับประเภทแบบทดสอบ (2) นักศึกษาทั้งสองกลุ่มมีเจตคติที่ดีต่อแบบทดสอบการพูดบนเว็บไซต์สำหรับภาษาอังกฤษเพื่อการท่องเที่ยวและความเห็นของทั้งสองกลุ่มไม่ต่างกันอย่างมีนัยสำคัญทางสถิติ (3) นักศึกษาที่มีความสามารถทางภาษาสูงใช้กลวิธีในการทำข้อสอบค่าความถี่สูงกว่า และประเภทของกลวิธีในการทำข้อสอบมากกว่านักศึกษาที่มีความสามารถทางภาษาต่ำกว่าในแบบทดสอบสามประเภทบนแบบทดสอบภาษาเฉพาะทางออนไลน์ แต่ความแตกต่างในด้านประเภทของกลวิธีในการทำข้อสอบมีไม่มากนัก ผลการวิเคราะห์ของงานวิจัยนี้สามารถนำไปประยุกต์ใช้ในการสร้างและวัดผลภาษาเฉพาะทางในบริบทอื่น โดยใช้เทคโนโลยีเป็นส่วนประกอบ

สาขาวิชา ภาษาอังกฤษเป็นภาษานานาชาติ ลายมือชื่อนิสิต .....

ปีการศึกษา ..2554

ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....

## 5187805820 : MAJOR ENGLISH AS AN INTERNATIONAL LANGUAGE  
 KEYWORDS: LANGUAGE FOR SPECIFIC PURPOSES/SPEAKING  
 ABILITY/ENGLISH FOR TOURISM/INTERNET ASSESSMENT/TEST  
 TAKERS' ATTITUDES/TEST TAKING STRATEGIES

MALINEE PHAIBOONNUGULKIJ : THE EFFECTS OF WEB-BASED  
 SPEAKING TEST IN ENGLISH FOR TOURISM (WBST-EFT) TASK  
 TYPES AND ENGLISH PROFICIENCY LEVELS ON STUDENTS'  
 SPEAKING PERFORMANCE AND AN INVESTIGATION OF THEIR  
 ATTITUDES AND TEST TAKING STRATEGIES. ADVISOR: PROF.  
 EMERITUS KANCHANA PRAPPHAL, Ph.D., 285 pp.

The objectives of this study were: (1) to examine the effects of the task types of the WBST-EFT and English proficiency levels on English for Tourism speaking performances in terms of language and content knowledge, and fluency; (2) to study the students' attitudes towards the WBST-EFT; and (3) to investigate and compare the types and frequency of strategies used by high and low proficiency students in doing the WBST-EFT. A total of 120 third year students who took English for Tourism II course at Nakhon Ratchasima Rajabhat University participated in this study. Stratified sampling was employed to classify the subjects into high and low proficiency groups. The research instruments were a needs analysis questionnaire, a web-based speaking test in English for Tourism (WBST-EFT), the attitudes towards the WBST-EFT online questionnaire, and a speaking test taking strategies coding scheme. Two-way ANOVA was conducted to investigate both the main and interaction effects between the proficiency levels and the task types on the test performances. Additionally, independent samples t-test was employed to compare the attitudes towards the WBST-EFT between the two proficiency groups. Descriptive statistics, frequency and percentage were employed to describe and compare the language for specific purposes (LSP) speaking components, students' views and their test taking strategies. Content analyses on speech performances and test taking strategies' verbal reports were carried out to obtain the in-depth information.

The findings were as follows. First, proficiency levels had a significant main effect on the LSP speaking performances. Additionally, there was a significant task types effect on vocabulary, language functions and content knowledge. More specifically, task types had greater impact on the low proficiency group than on the high proficiency group, particularly in the content knowledge component. Content analysis revealed prominent features of some speaking components associated with the task types. Second, the two proficiency groups' views towards the WBST-EFT were positive and not significantly different. Finally, the high proficiency students reported higher frequency and more types of strategies used than the low proficiency group across the three different task types in the online LSP test. However, the difference in types of strategies was not obvious. The results of the current study can be applied to other LSP technology-integrated tests.

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

Academic Year : 2011 Advisor's Signature .....

## ACKNOWLEDGEMENTS

I would like to express my gratitude to the Office of the Higher Education Commission, Thailand according to the Strategic Network Project for Production and Development of High Educational Institution Instructors in the Domestic Doctoral Degree Program for supporting my study. I am deeply indebted to Professor Emeritus Dr. Kanchana Prapphal, my advisor, for her valuable guidance, support and encouragement throughout this research. My gratitude also goes to all committee members: Assistant Professor Dr. Kulaporn Hiranburana, Assistant Professor Dr. Pavinee Thirakhupt, Assistant Professor Dr. Somsak Boonsathorn and Ajarn Dr. Sutthirak Sapsirin for their comments and suggestions.

My special thanks go to Associate Professor Dr. Pornpimol Senawong, Assistant Professor Dr. Thawascha Dechsubha, Assistant Professor Dr. Wiwat Puntai and Ajarn Dr. Wannana Soontornnaruerangsee for their advice on the instrument validation. I also thank Ajarn Onuma Lakarnchua from Chulalongkorn University Language Institute for her assistance in developing the test. I am grateful to Ajarn Pisit Puangmaliwan and Ajarn Bu-Nga Kanawarong from the English Program, Nakhon Ratchasima Rajabhat University for rating the test. I am thankful to Ajarn Piyachat Dhephasadin Na Ayudhaya for her great support in the academic writing. I also want to thank the English Program, the Faculty of Humanities and Social Sciences, Nakhon Ratchasima Rajabhat University for supporting me during my data collection process.

Finally, I would like to thank all of the EIL instructors and friends. Above all, I am truly grateful to my family, my uncle, my aunt, especially my mother, who always cheers me up during my hard time.

## CONTENTS

	Page
Abstract (Thai).....	iv
Abstract (English).....	v
Acknowledgements.....	vi
Contents.....	vii
List of Tables.....	x
List of Figures.....	xii
CHAPTER I INTRODUCTION.....	1
1.1 Background of the study.....	1
1.2 Research questions.....	2
1.3 Objectives of the study.....	3
1.4 Statement of hypotheses.....	3
1.5 Scope of the study.....	4
1.6 Limitations of the study.....	4
1.7 Definition of terms.....	5
1.8 Significance of the study.....	7
1.9 Overview of the study.....	8
CHAPTER II LITERATURE REVIEW.....	10
2.1 Language for specific purposes (LSP) ability.....	10
2.1.1 Definitions and components of LSP ability.....	10
2.1.2 Studies on LSP ability.....	14
2.2 Language for specific purposes (LSP) speaking tests.....	18
2.2.1 Standardized LSP speaking tests.....	19
2.2.2 English for Tourism II speaking test at Nakhon Ratchasima Rajabhat University.....	26
2.3 Web-based language testing.....	28
2.3.1 The use of web-based in language assessment.....	28
2.3.2 Web-based and computer-based speaking tests.....	30
2.3.3 Advantages and limitations of web-based language testing.....	36
2.4 The effects of speaking test tasks on speaking performances.....	38
2.5 Test takers' attitudes towards web-based/computer-based speaking tests.....	43
2.6 Strategies used in speaking tests.....	46
2.6.1 Taxonomy of strategies used in speaking tests.....	46
2.6.2 Studies on strategies used in speaking and oral tests.....	48
2.6.3 Studies on strategies used in web-based/computer-based speaking tests.....	57
2.7 Studies on research methodology.....	61
2.8 Summary.....	64
CHAPTER III RESEARCH METHODOLOGY.....	66
3.1 Research design and approach.....	66
3.2 Population and sample.....	66

	Page
3.3 Research stages .....	68
3.4 Research instruments .....	69
3.4.1 The Web-based Speaking Test in English for Tourism (WBST-EFT) .....	69
3.4.2 The Attitudes towards the WBST-EFT Online Questionnaire ...	80
3.4.3 The Speaking Test Taking Strategies Interview .....	82
3.5 Data collection .....	83
3.6 Rating for test scores.....	83
3.7 Data analysis .....	84
3.8 Summary .....	85
 CHAPTER IV RESULTS.....	 86
4.1 Research question 1 .....	86
4.1.1 The effects of proficiency levels and task types on the total LSP speaking scores .....	87
4.1.2 The effects of proficiency levels and task types on the LSP individual component scores .....	88
4.1.3 The effects of task types on the LSP individual component in each proficiency level.....	103
4.1.4 Content analysis of the LSP speaking performances from high and low proficiency groups .....	109
4.2 Research question 2 .....	153
4.2.1 Students' attitudes on the four aspects of the WBST-EFT .....	153
4.2.2 Students' attitudes on the individual aspect of the WBST-EFT .	154
4.2.3 Content analysis of the open-ended part .....	158
4.3 Research question 3 .....	161
4.3.1 Comparison of proficiency levels and reported strategies .....	162
4.3.1.1 Comparison of proficiency levels and total reported strategies .....	162
4.3.1.2 Comparison of proficiency levels and individual sub-category.....	163
4.3.2 Comparison of task types, proficiency levels and reported strategies .....	166
4.3.2.1 Comparison of task types, proficiency levels and total reported strategies .....	166
4.3.2.2 Comparison of task types, proficiency levels and individual sub-category.....	169
4.3.3 Content analysis from the verbal reports .....	176
4.4 Summary .....	191
 CHAPTER V DISCUSSIONS AND CONCLUSIONS .....	 192
5.1 Summary of the study .....	192
5.2 Discussions .....	195
5.2.1 Effects of the WBST-EFT task types and proficiency levels on speaking performances .....	195



	Page
5.2.2 Students' attitudes towards the WBST-EFT .....	205
5.2.3 Speaking test taking strategies used by the high and low proficiency students in doing the WBST-EFT .....	206
5.3 Conclusions .....	210
5.4 Implications of the study .....	212
5.4.1 Theoretical implications .....	212
5.4.2 Practical implications .....	213
5.5 Recommendations .....	214
REFERENCES .....	216
APPENDICES .....	225
APPENDIX A: A Needs Analysis Questionnaire .....	226
APPENDIX B: Web-based Speaking Test in English for Tourism (WBST-EFT) .....	231
APPENDIX C: The Rating Scale .....	241
APPENDIX D: The Attitudes towards the WBST-EFT Online Questionnaire .....	250
APPENDIX E: List of Speaking Test Taking Strategies, Speaking Test Taking Strategies Elicitation Instructions, and Interview Scripts and Questions .....	258
APPENDIX F: Samples of the Speech Performances and the Test Scores .....	263
APPENDIX G: The Speaking Test Taking Strategies Coding Scheme ..	271
BIOGRAPHY .....	285

## LIST OF TABLES

Table	Page
4.1: The effects of proficiency levels and task types on LSP speaking scores using 2x3 ANOVA .....	87
4.2: Descriptive statistics of the speaking scores of the six groups .....	88
4.3: The effects of proficiency levels and task types on pronunciation scores.....	89
4.4: Descriptive statistics of pronunciation scores.....	90
4.5: The effects of proficiency levels and task types on vocabulary scores .....	91
4.6: Scheffé post- hoc test on vocabulary scores .....	92
4.7: Descriptive statistics of vocabulary scores .....	92
4.8: The effects of proficiency levels and task types on grammar scores.....	93
4.9: Descriptive statistics of grammar scores .....	94
4.10: The effects of proficiency levels and task types on language functions scores .....	95
4.11: Scheffé post- hoc test on language functions scores.....	96
4.12: Descriptive statistics of language functions scores.....	96
4.13: The effects of proficiency levels and task types on cohesion scores.....	97
4.14: Descriptive statistics of cohesion scores.....	98
4.15: The effects of proficiency levels and task types on fluency scores .....	99
4.16: Descriptive statistics of fluency scores .....	100
4.17: The effects of proficiency levels and task types on content knowledge scores .....	101
4.18: Scheffé post- hoc test on content knowledge scores.....	102
4.19: Descriptive statistics of content knowledge scores.....	102
4.20: The summary table of the mean difference in each speaking component of the high proficiency students.....	104
4.21: Scheffé post- hoc test in high proficiency groups.....	105
4.22: The summary table of the mean difference in each speaking component of the low proficiency students.....	106
4.23: Scheffé post- hoc test in low proficiency groups.....	108
4.24: Mean differences of the students' attitudes on the four aspects of the WBST-EFT.....	154
4.25: Mean differences of the students' attitudes on the usefulness of the WBST-EFT.....	155
4.26: Mean differences of the students' attitudes on the appropriateness of time for preparation and response formulation of the WBST-EFT .....	156
4.27: Mean differences of the students' attitudes on the task difficulty of the WBST-EFT.....	157
4.28: Mean differences of the students' attitudes on the interface design of the WBST-EFT.....	157
4.29: Proficiency levels and total reported strategies .....	162
4.30: Sub-categories of Cognitive strategy .....	163
4.31: Sub-categories of Metacognitive strategy.....	164
4.32: Sub-categories of Communication strategy.....	165
4.33: Comparison of the task types, proficiency levels and total reported strategies .....	166

Table	Page
4.34: Sub-categories of Cognitive strategy used by high and low proficiency groups .....	169
4.35: Sub-categories of Metacognitive strategy used by high and low proficiency groups .....	171
4.36: Sub-categories of Communication strategy used by high and low proficiency groups .....	173

**LIST OF FIGURES**

Figure	Page
3.1: Group assignment with 2x3 factorial design using stratified sampling.....	67
3.2: The WBST-EFT framework .....	70

# CHAPTER I

## INTRODUCTION

### **1.1 Background of the study**

The sophistication of the world wide web has transformed testing format and administration to be available to a vast number of test takers, everywhere, with cost-effectiveness and decreased labor-intensiveness in both scoring and proctoring (Roever, 2001, Hamilton, Klein and Lorie, 2000, and Garcia Larboda, 2007a). The paradigm shift in the technology-integrated test format is notable in prominent English standardized tests, including the two leading tests, the Test of English as a Foreign Language internet-based test (TOEFL iBT) and International English Language Testing System computer-based test (IELTS CBT) (Garcia Larboda, 2007a and Alderson, 2009).

Despite the advantages over their preceding versions, CBT, particularly on the practicality issues regarding less reliance on technological expertise and logistic flexibility (Roever, 2001), web-based testing (WBT) has been purposively used in the present study for speaking skill assessment. Although WBT has been used in a vast number of testing areas, its use to assess speaking ability is limited (Garcia Larboda, 2007b), particularly in the language for specific purposes (LSP) domains; and thus, requires further investigations on the issues concerning factors affecting speaking test performances. According to Bachman and Palmer (1996), test tasks are considered one of the potential factors affecting language test performances. They have been potentially explored in the testing context (Turner and Uspher, 1995, Lumley and O'Sullivan, 2005, Teng, 2008, and Cumming et al., 2004); however, the results were varied. In this study, types of test tasks were obtained from the content of an English for Tourism course (EFT) which is an LSP course at Nakhon Ratchasima Rajabhat University. English for Tourism is one of the potential LSP areas in both business and educational context. The tourism enterprise was one of the major sources of income for Thailand that created approximately 6.7% of the kingdom's gross domestic product. This was the result of approximately 18.82 million tourists arriving in 2011 (Thailand Tourist Arrivals, 2011). For this reason, a large number of educational institutions offered English for Tourism courses to produce proficient English

speaking staff, particularly tour guides. Since tour guides are key individuals in various tourism enterprises who directly communicate with linguistically diverse tourists, English speaking skill is essential for their chosen career. Due to the increasing number of English for Tourism courses and the need to be able to evaluate LSP skills and abilities, there is a strong need for an assessment instrument that can precisely and accurately measure the speaking ability of these LSP majors.

Apart from the task types, test takers' characteristics have been noted as important factors affecting the test performances. These characteristics include examinees' target language proficiency levels, attitudes towards the innovative test (Kenyon and Malabonga, 2001, Norris, 2001, Kenyon, Malabonga and Carpenter, 2001 and Warschauer, 1996) and test taking strategies (Cohen, 1998, Song, 2005, Ting and Phan, 2008, Cabaysa and Baetiong, 2010 and Mendez Lopez, 2011). These topics were explored in a few studies with inconclusive findings. However, very limited studies have been conducted on the web-based speaking test (WBST) (Swain et al., 2009). Therefore, the issues concerning the effects of both task types and English proficiency levels on students' performances in LSP context, their attitudes and test taking strategies towards the innovative test require further investigation and they are the focus in the present study.

## **1.2 Research questions**

1. To what extent do the WBST-EFT task types and proficiency levels affect the speaking performances of high and low proficiency students in the English for Tourism course in terms of language and content knowledge, and the fluency of their speech performances?
2. What are the students' attitudes towards the WBST-EFT?
3. Are there any differences in the types and frequency of speaking test taking strategies used by high and low proficiency students in doing the WBST-EFT?

### 1.3 Objectives of the study

1. To study the effects of the task types of the WBST-EFT and English proficiency levels on English for Tourism speaking performances in terms of language and content knowledge, and fluency.
2. To examine the students' attitudes towards the WBST-EFT.
3. To investigate and compare the types and frequency of strategies used by high and low proficiency students in doing the WBST-EFT.

### 1.4 Statement of hypotheses

1. There is no significant difference between the mean score of the high proficiency students and that of the low proficiency students at the .05 level.

a.  $H_0: \bar{X}_H = \bar{X}_L$

2. There are no significant differences at the .05 level in the three task types performed by the two proficiency groups.

b.  $H_0: \bar{X}_{T1} = \bar{X}_{T2} = \bar{X}_{T3}$

3. There are no significant interaction effects at the .05 level between the two proficiency groups and the three task types.

c.  $H_0: \bar{X}_{H.T1} = \bar{X}_{H.T2}$

$H_0 \bar{X}_{H.T1} = \bar{X}_{H.T3}$

$H_0 \bar{X}_{H.T2} = \bar{X}_{H.T3}$

$H_0: \bar{X}_{L.T1} = \bar{X}_{L.T2}$

$H_0: \bar{X}_{L.T1} = \bar{X}_{L.T3}$

$H_0: \bar{X}_{L.T2} = \bar{X}_{L.T3}$

4. There is no significant difference in the attitudes towards the WBST-EFT in the two proficiency groups at the .05 level.

d.  $H_0: \bar{X}_H = \bar{X}_L$

### **1.5 Scope of the study**

1. The population was 230 students who took English for Tourism II course at Nakhon Rachasima Rajabhat University (NRRU) in the second semester of the academic year 2010. The sample group was 120 students who were randomly selected and classified into six groups (three high proficiency groups and three low proficiency groups) using the stratified sampling technique.

2. Web-based Speaking Test in English for Tourism (WBST-EFT) is the final achievement test for English for Tourism II course at NRRU. The test content focuses only on three task types that are related to the course syllabus and the results from a needs analysis questionnaire. The task types include presenting tourism-related information, giving polite suggestions and responding to tourists' enquiries and complaints.

3. Students' speaking performance in English for tourism includes some of Douglas's LSP ability (2000) and Fulcher's speaking ability (2003a): knowledge of pronunciation, vocabulary, grammar, language functions, cohesion, fluency and content knowledge.

4. Speaking test taking strategies in this study focus on some of the strategies used in the internet-based test as proposed by Swain et al. (2009), namely, Communication, Cognitive and Metacognitive strategies.

### **1.6 Limitations of the study**

1. The sample group of this study was the students at NRRU, so the generalizability of the findings must be interpreted with caution.



2. The task types were derived primarily from the course content analysis and these types might not cover all the actual task types performed by professional tour guides. Therefore, the inference of the test scores must be cautiously interpreted.

### **1.7 Definition of terms**

**Web-based Speaking Test in English for Tourism (WBST-EFT)** is a semi-direct speaking test that includes three speaking task types. The test tasks are created on Moodle Version 1.9.5 which is a freeware online template and program that is currently used at NRRU and the test is administered via the internet system at the computer laboratory. The test is used as the final achievement test to assess the students' overall speaking proficiency and the knowledge of the topics that have been taught in English for Tourism II course. The students are provided with headsets and asked to respond to the test tasks by speaking through microphones. Their speaking performances are recorded in the audio files and saved in the database. The performances are rated later by two raters.

**Task types of the Web-based Speaking Test in English for Tourism** refer to the three task types: presenting tourism-related information, giving polite suggestions and responding to tourists' enquiries and complaints. Types of tasks are derived from the English for Tourism II course content and objective analysis, and from the needs analysis questionnaire distributed to the subject specialists in tourism-related fields. All task types are taught in the course.

**Speaking performance** refers to the ability to use language knowledge and content knowledge taught in English for Tourism II course to respond to the three task types of the WBST-EFT. This includes the ability to produce fluent responses. The speaking performance is assessed on accuracy, range, complexity and appropriateness of the language production as described in the rating scale. The average score from two raters in each test task and each component is the representation of the speaking performance.

**Language knowledge** refers to knowledge of pronunciation, vocabulary, grammar, language functions and cohesion. The language knowledge is measured by the average score derived from the rating scale from two raters from the summation of the five language components for each test task.

**Knowledge of pronunciation** is the ability to use sound, stress and intonation to convey the intended meaning of an utterance. It is measured by the degree of accuracy to pronounce words that conform to standard varieties of English. It also includes the effective use and degree of intelligibility of stress to emphasize particular words and use of intonation to convey speech functions. The average score from two raters is the representation of this knowledge.

**Knowledge of vocabulary** is the ability to use both generic and tourism-related technical terms to respond to the test tasks. It is measured by the accuracy and range of the vocabulary employed in the responses. The knowledge of the vocabulary is demonstrated through the average score from two raters.

**Knowledge of grammar** is the ability to use standard English grammatical structures and rules to produce comprehensible performances. It includes the use of specific language patterns to construct appropriate responses to the test tasks. It is assessed by the accuracy, range, complexity and appropriateness of the structures in the speech produced. The average score from two raters is the representation of this knowledge.

**Knowledge of language functions** means the ability to interpret and formulate appropriate and logical speech. It includes the use of ideational, manipulative, heuristic and imaginative functions to respond to the test tasks. Knowledge of language functions is measured by the appropriateness of the speech produced in terms of the meanings, task requirements, and language use setting. The average score from two raters is the representation of this knowledge.

**Knowledge of cohesion** is the ability to combine phrases and sentences in a meaningful way, which can be seen from the use of cohesive devices in the responses. It is measured by the accuracy and range of cohesive markers in the test

performances. Knowledge of cohesion is demonstrated through the average score from two raters.

**Fluency** is the general quantity and tempo of language production. It is the ability to use the tempo and pauses in language production to maintain paces of the responses. It is assessed by the appropriate use of both the tempo and pauses in the responses. Fluency of speech is demonstrated through the average score from two raters.

**Content knowledge** is the ability to present tourism-related content knowledge taught in English for Tourism II. It is measured by the accuracy and completion of the information given by the students to respond to the test tasks. Content knowledge is demonstrated through the average score from two raters.

**Attitudes** mean the students' perceptions towards the Web-based Speaking Test in English for Tourism (WBST-EFT) concerning the overall usefulness of the test, appropriateness of time for preparation and response formulation, task difficulty and the interface design. Attitudes are derived from the content and statistical analysis of the students' attitudes towards the WBST-EFT online questionnaire.

**Speaking test taking strategies** refer to types of conscious thoughts and reported behaviors from the students while they are attempting each test task. Speaking test taking strategies are classified into three main types: Communication strategy (achievement and avoidance), Cognitive strategy (selecting, comprehending, storing memory and retrieval) and Metacognitive strategy (goal setting, organizing, planning and evaluating). Types and the frequency of the reported strategies are classified and calculated.

## **1.8 Significance of the study**

1. For theoretical contributions, the research study can provide information on the effects of particular factors on the language for specific purposes (LSP) test performances with web-based speaking assessment. The findings can also reveal whether there is interaction effect between task types and English proficiency levels on the LSP speaking performances in technology-integrated test. The study is

expected to yield the empirical evidence regarding construct and content of English for Tourism when using the internet to assess speaking ability. The information on students' attitudes towards this type of testing and their test taking strategies can be obtained as part of the validation study. The finding can provide information on the particular types of strategies used by certain group of students and in particular tasks in the new mode of test delivery.

2. For pedagogical contributions, the findings can provide information for the university about the usefulness of the WBT for assessing students' speaking performances for large classes. The results of the study can be used in course development for teachers and for test development in other LSP courses at NRRU. The strategies reported by the two proficiency groups will be beneficial in providing information that can be used in developing speaking skill for English for Tourism II course.

### **1.9 Overview of the study**

This chapter presents the background of the study. The research questions and objectives in relation to the problems are covered. It also includes hypotheses, scope, limitations, definition of terms, and significance of the study.

Chapter 2 provides the review of literature covering language for specific purposes (LSP) speaking ability and tests, web-based language testing, the effects of speaking test tasks on speaking performances, test takers' attitudes towards the web-based and computer-based tests and test taking strategies.

Chapter 3 focuses on research methodology. It includes population and sample, research design and approach, research instruments, data collection and data analysis.

Chapter 4 provides the results of the main study in relation to the three research questions. Quantitative and qualitative analyses are presented.

Chapter 5 summarizes the research findings, and the discussions of each research question are presented. Conclusions of the findings, implications and recommendations are provided in this chapter.

## **CHAPTER II**

### **LITERATURE REVIEW**

Review of related literature includes language for specific purposes (LSP) speaking ability, discussion of the studies on LSP language ability, LSP speaking tests, web-based language testing (WBT), advantages and limitations of web-based language tests, the effects of speaking test tasks on speaking performances, test takers' attitudes towards web-based and computer-based tests and their test taking strategies.

#### **2.1 Language for specific purposes (LSP) speaking ability**

##### **2.1.1 Definitions and components of language for specific purposes (LSP) ability**

Language for specific purposes (hereafter LSP) is the significant sub-field in testing in which test takers' performances can be used as the inference for precise language ability in a particular target language use situation. LSP ability is defined by Douglas (2000) as

Specific purpose language ability results from the interaction between specific purpose background knowledge and language ability, by means of strategic competence engaged by specific purpose input in the form of test method characteristics (Douglas, 2000:40).

The author proposes one of the most prominent frameworks for LSP ability which can be adapted into four skills including speaking ability. This framework is based on Bachman and Palmer's (1996) model of communicative language ability with the modification on the strategic competence component and the emphasis on the notion of background knowledge.

In specific purposes context, the relationship between language ability and specific background knowledge is one of the key features for language for specific

purposes. The LSP ability model comprises of language knowledge, strategic competence and background knowledge.

Language knowledge incorporates grammatical knowledge, textual knowledge, functional knowledge and sociolinguistic knowledge. This knowledge deals with the process of language production and its appropriate use in the context.

Grammatical knowledge deals with the combination of the linguistic elements including vocabulary, morphology and syntax, and phonology to produce and comprehend the standard and the grammatical form of language.

Textual knowledge enables the language users to combine these linguistic components into the larger construction which is the unit of language or discourse. It involves knowledge of cohesion (the study of the relationship among sentences) and rhetorical or conversational organization (the organizational development of a text) in written and spoken discourse.

Functional knowledge is made up of knowledge of ideational functions, knowledge of manipulative functions, knowledge of heuristic functions and knowledge of imaginative functions (Douglas, 2000). This classification is based on pragmatic knowledge of Bachman and Palmer (1996). From this classification, “functional knowledge enables us to interpret the relationships between utterances or sentences and texts and the intentions of language users” (Bachman and Palmer, 1996: 69). The first category of functional knowledge is knowledge of ideational functions that enables us to express or interpret meaning from our real world’s experience. This knowledge deals with the use of language to describe, classify, explain knowledge and ideas, and to express feelings. The second category is knowledge of manipulative functions that enables us in using language to affect the world around us. This functional knowledge comprises three functions: instrumental functions (getting people to do things), regulatory functions (controlling what people do), and interpersonal functions (making, maintaining and changing interpersonal relationships). The third category is knowledge of heuristic functions that enables us to use the language to extend our world’s knowledge. The last category is knowledge

of imaginative functions that enables to use the language to create an imaginary world, and to extend our environment for humorous and aesthetic purposes.

The classification of functional knowledge is from language functions' framework proposed by Halliday (1973, 1976 cited in Bachman, 1990). In this framework, language functions include four types of macro-functions: ideational, manipulative, heuristic and imaginative. Ideational function is employed in expressing meanings of our experience of the real world; and it includes the use of language in expressing, exchanging and interpreting meaning, ideas, knowledge and feelings. On the other hand, manipulative functions are used to "affect the world around us" (Bachman, 1990:93), and three types of functions are included: instrumental (dealing with how to get things done), regulatory (dealing with the control of others' behaviors, formulating and stating rules, laws and norms of behavior), and interpersonal (forming, maintaining or changing interpersonal relationships). Heuristic function involves the use of language to "extend the knowledge of the world around us" (Bachman, 1990:93) such as the language of teaching and learning, problem solving and memorization of the information. The last type of language functions, imaginative function enables us to create and extend our world for creativity and enjoyment purposes.

Sociolinguistic knowledge is the production, use and interpretation of the language with the consideration of a particular context. This conventional knowledge includes the appropriate formulation and comprehension of language for a specific region and social group (dialects), certain stylistic of language (register), in native-like manner (naturalness) and in different cultures and figure of speech (cultural reference and figure of speech).

Strategic competence refers to the metacognitive strategies or higher order thinking and communication strategies which are hierarchically employed by language users. These strategies relate the interaction of language users' internal language ability (language knowledge and background knowledge) to external context in planning a communicative response. While higher order thinking strategies will be used in dealing with general features of context that is not involving language use; the



communication strategies will function in more specific context and directly link with the language use respectively. In this framework communication strategies are the integral competence in LSP testing context. The strategies are discourse assessment, communicative goal setting, linguistic planning, and control of linguistic execution.

Discourse assessment deals with analyzing characteristics of the specific purpose situation and deciding on the selection of relevant discourse domain (internal interpretation) of the context in the communicative situation which includes the setting, participants, purpose, message form and content, tone, language, norms of interaction and genre. This strategy is the continual and reiterative process that directly affects goal setting, planning and execution employed by language users. It is also used to adapt the discourse to facilitate the incomprehensible message or to prevent the breakdown of the conversation.

Setting communicative goal involves the decision making on selective use of discourse domain with the consideration of communicative goal. In a testing situation, test takers may have their own goal to respond effectively and rapidly to a semi-direct speaking task. To achieve the goal, they are required to identify the task and choose the appropriate discourse domain in a limited time.

Linguistic planning is the sub-sequent strategy of goal setting. It deals with the allocation of appropriate and pertinent background knowledge and language knowledge (grammatical, textual, function and sociolinguistic knowledge) to achieve the communicative goal. In some cases, certain strategies: avoidance, paraphrase, translation, support requirement and gesture will be used to reach the goal when the background knowledge and language knowledge are not available.

Control of linguistic execution is the final strategy to carry out the response. Appropriate background knowledge and language knowledge are retrieved, systematically arranged and related in the language production or comprehension in appropriate means. Due to dynamic and complicated nature of context, language users can possibly misinterpret the unfamiliar context resulting in unintelligibility and miscommunication. Another difficulty is the inadequacy of language ability to formulate the response. These problems can be overcome with the sufficient provision

of signals as “contextualized cues” e.g. change in voice tone, pitch, tempo, rhythm, gaze and facial expression to create mutual understanding among language users e.g. in test method characteristics, test prompt and input data (Gumperz, 1976 cited in Douglas, 2000:44).

Background knowledge is the central issue that marks distinctive characteristics of LSP ability. Individuals relate this long term memory knowledge based on previous experience with the present input to predict the upcoming events and make a decision. In the testing context, test takers retrieve pertinent background knowledge and relate it to language knowledge to interpret the communicative situation and respond to the test tasks that resemble the target language use situation. Along the process, communication strategies serve as the mediator to facilitate the interaction between these language ability components.

## **2.1.2 Studies on language for specific purposes (LSP) ability**

### **2.1.2.1 Studies on strategic competence**

Although both influential models of language ability from Bachman and Palmer (1996) and Douglas (2000) combine strategic competence in the language ability, Elder (2001) points out the problematic issue on this inclusion. She reported the ‘mismatch’ of the test takers’ LSP communicative ability between linguistic competence and non-language ability, for example strategic competence and teaching skills. Her finding from a Rasch analysis of scores obtained from test takers of Language Proficiency Test for Teachers of Italian and Japanese as a Foreign Language in Australia (thereafter LPTT), a direct speaking test, revealed the performance scores’ incompatibility. This performance-based test required the test takers to deliver the lesson for elementary students. 7.5% of the test takers who scored high on linguistic components failed to achieve the score on strategic competence while the reverse was true for test takers whose strategic competence score was high. High linguistic proficient test takers outperformed in describing the classroom activity, folding papers into animal shapes, with the use of sophisticated language construction. However, they were rated as insensitive to audience by using too complicated structures, hence, would not be pertinent to young learners’ language

proficiency. In reverse, low language competent test takers employed accommodation strategies with the use of simple construction and non-verbal language including hand movement to demonstrate the lesson. This problem as proposed by the author could be entangled by the separation of the linguistic ability from non-language competence and use the latter ability as additional consideration in scoring. This solution could be employed in the situation that the sensitivity to context can substitute the inadequate linguistic competence.

### **2.1.2.2 Studies on background knowledge**

The prominent feature of LSP test is the inclusion of background knowledge in test construct as asserted by Douglas (2000) saying that

The very essence of specific purposes language tests is that they require the test takers to engage themselves authentically in the test tasks that are demonstrably related to the target language use situation, and, therefore, relevant background knowledge will necessarily be called upon in the interpretation of the communicative situation and in the formulation of a response (Douglas, 2000:39).

However, the inclusion of the field specific content knowledge is questioned by Wu and Stansfield (2001) regarding their work on the Listening Summary Translation Exam in Taiwanese (LSTE/T), which is in law-oriented context. They argue that the language in the test tasks is the key feature that creates specificity of the test along with the authenticity of the test tasks.

Apart from the controversial issue concerning the inclusion of background knowledge in test construct, a number of studies investigated the effects of this field specific knowledge on language test performance. However, the findings of these studies remain inconclusive (Clapham, 1996 cited in Douglas, 2000 and Krekeler, 2006).

Clapham (1996, cited in Douglas, 2000) studied ten reading sub-tests performance of the International English Language Testing System (IELTS) with three proficiency groups. She found that students achieved higher scores on the

reading test in their own subject areas than on the general topics. Considering the levels of education, there was no significant subject areas effect on the undergraduate students, but the effect on the subject areas was found in the postgraduate group. It was possible that as the reading texts become more subject specific, background knowledge would have the stronger effect on the test scores. The finding also revealed that there was a highly significant effect of subject areas on the students with the score more than 60% on the grammar test. In contrast, the students with the score lower than 60% did not benefit from their background knowledge. Therefore, the levels of language knowledge, particularly on grammar influenced the effect of background knowledge on test performance.

The finding also indicated that among the three proficiency groups, the intermediate students did benefit most from their background knowledge. In other words, there was a strong effect of background knowledge on their reading test performance. It was explained by Clapham that the high proficient students could take the maximum advantage of their linguistic knowledge; thus, they did not need to rely on their background knowledge. As for the low level students, they were too concerned with bottom-up skills until they did not make use of their background knowledge. This notion that background knowledge varies in its effect on the test performance according to the levels of proficiency is later called 'the two thresholds effect'. However, Clapham's findings are requested for further clarification from Alderson (2000).

The notion of two thresholds effect on reading performance has also been reinvestigated in the study of Krekeler (2006) with some contradictory results to Clapham's findings. More than 500 subjects participated in this study. Two discipline-related business and technical texts were selected and the C-test scores were used as a measure of L2 proficiency. In general, there was a strong effect of background knowledge on reading performance. The finding revealed that students performed better in their own subject area topics than without them regardless of their L2 proficiency levels. The interaction effect between background knowledge and L2 proficiency levels was limited. The majority of students were able to take advantage of their background knowledge. The finding in technical-related texts was

contradictory to the two thresholds hypothesis in that the medium level students profited least from their background knowledge. However, the finding that subsumed the two thresholds hypothesis was only found in reading tests on business topics. The medium level language proficiency group did benefit most from their background knowledge whereas the low level group achieved similar scores on the texts in and outside their subject areas. In the high level group, the students did not depend mostly on their background knowledge. The author concluded that the notion of linguistic threshold should not be used to interpret the result of language for academic purposes testing when the text posed a low degree of specificity. In addition, it should be assumed that test performance will be affected from background knowledge or from the lack of it. It was suggested for test users to be aware of the influence of background knowledge on reading tests' results and the degree of effect could not be predicted from L2 proficiency levels.

### **2.1.2.3 Studies on specificity of language for specific purposes (LSP) ability**

Another issue mentioned by a number of scholars is the overlapping line between language for specific purposes testing and language for general purposes testing. Davies (2001) asserts that LSP testing still lacks the theoretical basis. The distinctive characteristics of LSP tests are on communicative nature and derived from practical requirement. From its nature, LSP tests can only be judged from the pragmatic approach. Similarly, Cumming (2001) claims that there were no precise criteria that could distinguish LSP from general purposes (GP) based on his investigation on EFL/ESL composition assessment from international instructors. He could not gather the systematic evaluation data in which the assessment criteria and methods varied in contexts, instructors themselves and the methods. The way in which these international instructors distinguished LSP from GP was through their perception on the course curricula which consequently affected their assessment. LSP assessment consisted of more specific ranges of criteria derived from needs analysis and the expectation on students' achievement was limited. The emphasis of the criteria was on the form of written text. In contrast, instructors employed numerous scoring methods and wider ranges of grading criteria. The focal point of GP

assessment was perceived through both individual students and their progress. Additionally, Elder (2001) mentions on the need for LSP job-related assessment. She reports on the finding conducted in 1993 of the poor correlation from the standardized test score, IELTS, with the teachers' performance in a training course. She points out the problem of precise criteria on distinguishing specific features of teachers' discourse due to the immeasurable pedagogical contexts. This includes the judgement on specific features of the test tasks and contents in the particular setting. Wu and Stansfield (2001) also state on the need for the language testers' support to pinpoint the distinctive characteristics of LSP test.

On the other hand, Douglas (2000) mentions that LSP testing is a sub-category of language testing. The content and test taking method of LSP tests are resulted from the target language use analysis (TLU) (Douglas, 2000:19). He states that all tests are developed for some purposes and they will fit in the particular point of the continuum of specificity. LSP tests therefore pose certain and precise characteristics used by people in the fields which include specific pronunciation, vocabulary, word meaning, and sentence structures that people who are not in the field do not have thorough understanding. He proposes that certain characteristics of the tests that are shared by numerous target language use situations should be focused to entangle the problem of specificity. He also argues that background knowledge and language knowledge cannot be separated. The separation of these language ability components might be misleading to our present understanding of the nature of human cognition.

LSP testing is considered controversial in terms of the distinctive features from GP testing. Therefore, the issues of the inseparability between language knowledge and strategic competence of LSP testing, and the combination of field specific content knowledge in specific language ability require additional investigations.

## **2.2 Language for specific purposes (LSP) speaking tests**

LSP tests pose certain characteristics when taken into account the context of use. The following tests are selected in this present study to exemplify the combination of features of language ability and the content of the LSP tests.

## **2.2.1 Standardized language for specific purposes (LSP) speaking tests**

### **2.2.1.1 The Japanese Language Test for Tour Guides**

This face-to-face interview oral proficiency test is developed by the NLLIA Language Testing Center at the University of Melbourne, Australia. The purposes of the test are to provide the indication of candidates' proficiency levels in Tourism and related fields to the test users and as a selection process of the applicants to the Japanese tour guide training course. Test takers are experienced and non-experienced tour guides. The test format comprises of six sections derived from job analysis of the tour guide' role and lasts 30 minutes in total. The test takers are required to take the role of a tour guide whereas the interlocutor acts as a Japanese tourist in all five sections with the exception to the first section. Details of the test content are described below.

#### Section one: Warm-up

The first section aims to familiarize the applicants with six easy-short questions related to their personal experience. They are informed that this part will not be assessed and included in their scores.

#### Section two: Optional tours

This section deals with helping clients to select an optional tour. Applicants thus will be required to demonstrate their language functions regarding making recommendations, giving suggestions, advice and opinions and selling the tour.

#### Section three: Problem-solving

This section aims to assess applicants' ability to use the language to respond to the customers' complaints. The functions of language are showing sympathy, giving reasons, and showing agreements.

#### Section four: Cultural presentation

Applicants are assessed on their language ability to describe the features of Australian culture that meet with Japanese tourists' interests and pertinent to their

experience. Language functions to be assessed are explaining factual information and experience, making comparison, demonstrating, providing reasons and showing opinions.

#### Section five: Giving instructions or information

This section requires the applicants to provide information about facilities and give instructions. The language functions involve explaining facts, giving instructions and advice, providing sequential information, and telling where things are.

#### Section six: Explaining itinerary and tourist sites

Applicants are assessed on their language ability to explain a day tour program and describing tourist sites. The language functions in this section comprise giving time information, describing specific characteristics, telling where things are and providing time information and duration.

The scoring criteria are based on applicants' linguistic skill and task fulfillment. Linguistic criteria cover grammatical resources and expression, pronunciation, fluency, use of vocabulary, polite forms, and comprehension. Each language component is marked from 1 to 6 rating scale on descriptive band scale. The sum of the score is derived by adding and averaging the overall scores based on the criteria. Task fulfillment criterion is also based on the 1 to 6 rating scale with the use of band scale descriptions. The scores are given on applicants' enthusiasm, empathy, making something sound interesting, persuasiveness and awareness of interlocutor's needs or desire. All of these criteria are considered essential qualities in tour guide communication (Brown, 1995).

#### **2.2.1.2 The International Legal English Certificate (ILEC)**

The ILEC is developed by Cambridge University ESOL Examinations and aims to be used as the indication of lawyers and legal-related professions about their English proficiency levels in communicating with clients and international law-related staff. The test is based on the Common European Framework of Reference for Languages (CEFR) from 'Vantage' level (B2) to 'Effective Operational Proficiency'



level (C1) and lasts under 3.50 hours. There are four parts in the test: reading, writing, listening and speaking. Each skill is accounted for 25%. Due to the focus of this study on LSP speaking test, only the details of the speaking test will be emphasized.

The speaking sub-test comprises of four parts and lasts for 16 minutes. It is an interactional face-to-face interview with pairs of candidates and examiners. One of the examiners is both interlocutor and an assessor who will give the global assessment score for the candidate while the other acts as assessor and will not join the interaction. This person will give separate score for the candidates based on the Cambridge ESOL Common Scale for Speaking criteria.

Task types range from short answers, 1 minute long turn, interaction questions and answers between the candidates and a three-way discussion.

In part one 'Interview', the time is 2 minutes. This part is the questions and answers to the interlocutor. The candidates are assessed on the ability to respond to the interlocutor's questions and provide additional details of their answer to legal-related topics. This is followed by the 'Long turn' part which takes 7 minutes to complete. The focus is on candidates' ability to maintain a 'Long turn', giving information and opinions through discourse management. The candidates are required to select one topic out of the two and the written prompt will be given respectively. The candidate has 1 minute to prepare the topic and 1 minute to talk about their topic without any interruptions. The second candidate will ask questions about the topic. Then, they will take turn to complete the task. Part three 'Collaborative task' lasts for 4 minutes. The emphasis of this part is on candidates' ability to collaborate and negotiate in the conversation through an appropriate initiation and responding. Candidates are provided with oral and written instructions of the task in which they are required to discuss and negotiate to complete the task. The final part 'Discussion' takes 3 minutes to complete. This part is the follow-up from Part three in which candidates are required to discuss about the topic through the questions initiated by the interlocutor. The candidates are assessed on their ability to participate in the discussion.

Each candidate is assessed on his/her individual performance. The scoring criteria are based on grammar and vocabulary, discourse management, pronunciation and interactive communication. The first aspect of criteria concerns with accurate use of grammatical forms and vocabulary. This includes the range and appropriacy of structures and vocabulary to complete the task. The second criterion, discourse management, deals with the ability to connect the utterances into monologues and logically form the dialogue to develop themes and arguments (coherence and relevance). The ability to produce the appropriate length of discourse (extent) is included. The third criterion, pronunciation, concerns with stress and rhythm, intonation and effective articulation of each sound to produce comprehensible utterances. The last criterion, interactive communication, emphasizes on the ability to participate in discourse development through the discussion without hesitation (initiating and responding, hesitation). The sensitivity to the interactants to maintain the conversation is also focused (turn-taking).

The score will be reported based on the Cambridge ESOL Common Scale for Speaking criteria in which the passing grades are B2 'Pass', and C1 'Pass with Merit'. Reports of candidates' performance will be focused on each skill (ILEC Handbook, 2008).

### **2.2.1.3 SPEAK/TEACH Test**

The test is developed by the Graduate College Program, Iowa State University of Science and Technology (ISU). This diagnostic oral proficiency test is intended to assess the international teaching assistants (ITAs) English communication skill in daily life and in the classroom context in the U.S.; so that they can be assigned into the proper teaching duties for undergraduate students. The SPEAK/TEACH test which first administered in August 2009 comprises of two sections and lasts about 25 minutes. Candidates are required to arrive 1 hour before the test appointment for preparation and their performances from both sections are recorded. During the test preparation session, candidates are provided with topics related to their academic fields. Textbook about each topic is given and they can prepare their 5 minutes lesson

for the TEACH test in a quiet room. They can take notes on the papers provided in the room; however, dictionaries and other textbooks are not allowed.

The candidates start with the first face-to-face interview, the SPEAK test, which takes about 10 minutes. There are three parts in this test. The first part is the warm-up that aims to ease the candidates and they are required to answer some informal questions. This is followed by a 5-minute impromptu speaking test. Candidates are expected to provide immediate responses to the questions in two different topics including some follow-up questions. Topics covered in this part are talking about experience and events, or giving ideas and opinions and shorter questions about describing things with object or picture help. The last part, the role play of a specific daily life situation, requires the candidates to take the role as stated on the card prompt in describing a particular situation, such as shopping and making an appointment. They have 1 minute to prepare for the conversation and they are required to speak up to 2 minutes.

The second section, TEACH, concerns with LSP in the teaching context that requires candidates to take the role of the teacher teaching a particular subject related to their academic field. This section follows the SPEAK test right away and takes about 15 minutes to complete. The candidates have 2 minutes to prepare for their lesson by writing on the board or drawing diagrams. They teach the class for 5 minutes and the remaining 3 minutes are devoted for questions and answers after their teaching lesson. The audiences in the class are two to three raters, a proctor, and at least one undergraduate. They ask questions related to the contents on the presentation or typical class procedures. Candidates are exempted from the SPEAK test if they pose high listening and speaking skills from the standardized tests e.g. IELTS, TOEFL, TOEFL iBT. They will begin the short warm-up and proceed to the second test.

The scoring criteria are based on the listening and speaking abilities in five aspects. The first aspect is 'Overall language effectiveness and comprehensibility' which includes volume, fluency, effective use of grammar and vocabulary, pronunciation, rhythm and intonation, and appropriacy of speed in speaking. The

second aspect is ‘Listening ability’ that candidates can comprehend students’ questions. The third aspect deals with ‘Question handling and responding’ which is the ability to answer questions effectively, flexibly and thoughtfully while showing who is in charge. The fourth aspect ‘Communication skills’ includes logical and sequential explanation development, clarity of expression, use of supporting evidence, eye contact, use of chalkboard and enthusiasm/presence. The last aspect concerns with ‘Cultural ability’ which includes familiarity with cultural code, especially in the U.S. context, appropriacy of non-verbal behavior and rapport with class.

The candidates’ performance is rated based on the score 0-300 on seven bands with combined score of both SPEAK and TEACH tests. In general, two raters provide the score on both tests; however, in the case of 30 points discrepancy the third rater is required. There are four levels on the test: ‘Fully certified’ (level 1), ‘Certified with restriction’ (level 2 and 3), and ‘Not certified’. The teaching assistant is required to get level 1 certification. In the case of lower level, they are expected to take English 180 at the university. The score is reported on the ISU website for SPEAK/TEACH Tests (Online: <http://www.grad-college.iastate.edu/speakteach/>).

#### **2.2.1.4 The Speaking Test by Business Language Testing Service (BULATS)**

The test is developed by BULATS, the University Cambridge ESOL Examinations. The purpose of the test is to assess speaking skill in business-oriented workplaces. It is offered in English, French, German and Spanish. Candidates are employees who use a foreign language for their professions and students who study in this field where English is an important means. The BULATS Speaking Test is a separated test that can be taken on its own or with BULATS tests, for example the standardized version that included all four skills. It is a face-to-face oral interview for individual candidate with one oral examiner. Candidates’ performance is recorded and marked by other assessors. The test comprises of three parts and each part lasts for 4 minutes. The total time taken to finish the test is 12 minutes. The details of the test are given below.

### Part one: Interview

Candidates are required to answer the questions from an examiner about their personal information, work, studies and interests.

### Part two: Presentation

Candidates are provided with three topics and they are required to select one topic. They have 1 minute to prepare for the presentation. After the end of the presentation, the examiner asks questions regarding the presentation.

### Part three: Information Exchange and Discussion

Candidates have 1 minute to read the role-play situation on the given sheet. They ask the examiner questions to get the information that leads to the discussion of the related topics.

### Scoring criteria

Candidates' performance is assessed by the ALTE rating scales developed by the University of Cambridge ESOL Examinations under the CEFR framework. The band ranges from 'Beginner' (0) to 'Very Advanced' (6). Candidates are assessed on their abilities to use correct words and grammar (accuracy), range of vocabulary and sentence structures (range of language), pronunciation of words (stress and intonation), organization of their utterances (discourse management), participation in the conversation ( interactive communication).

### Score report

Candidates can get the score report as Test Report Form (TRF) in which their proficiency level is stated on 'high', 'mid' and 'low' of the CEFR/ALTE levels of the overall band. The reverse page is the information of the ALTE 'Can-do statement' for each level to facilitate the candidates in score interpretation. There are two types of reports which are group, organization, or test user report and candidate report (BULATS Test Specification: A Guide for Clients, 2009. pp.10-12).

### **2.2.2 English for Tourism II speaking test at Nakhon Ratchasima Rajabhat University**

The English for Tourism II speaking test is the final achievement test for the course in which test tasks and construct are primarily derived from the analysis of course description and objectives.

English for Tourism II is one of the English for specific purposes (ESP) courses offered in every second semester at Nakhon Ratchasima Rajabhat University (NRRU) as a three-credit compulsory course for third year Business English- and Tourism Industry-majored students and as an elective course for other majors. English for Tourism I is a pre-requisite for this course. The course content is related to “a study of English technical terms necessary for describing domestic tourist attractions and their histories in the central region. Studying the language for operating domestic trips for foreigners is focused. Organizing field trips to actual sites is recommended” (Curriculum for Bachelor of Arts in English, 2008:36). The objectives of the course are “to acknowledge students about history, culture and attractions in central region of Thailand and to practice using technical terms and useful expressions in variety of situations in tour operation context”. Students are expected to meet the following qualifications to pass the course:

1. Students should be able to effectively (accurately and meaningfully) use the technical terms and expressions in tourism context.
2. Students should be able to explain and present tour itinerary, attractions, brief history and culture.
3. Students should be able to use polite requests and respond to enquiries and complaints.

Although the course focuses on all four skills of the English language, the primary concern is in speaking which takes 60% of the total scores in the course (20% is contributed to the midterm test and the rest 40% is for the final test). The course grade will be used as part of the application for bronze card tour guide license for Tourism Industry-majored students. Therefore, parts of the course contents will

conform to those of the tour guide training curriculum offered by the Ministry of Tourism and Sports (Tourism Authority of Thailand Act, 1979). The details on the part of the foreign language instruction and assessment are as follows.

1. Practice one of the foreign languages by accurately giving the information on history, arts, culture, festivals and attractions with the consideration of cultural awareness of the native language.
2. Practice daily conversation.
3. Practice the expressions used by professional tour guides e.g. tour operation, giving polite suggestions and description of the tourist attractions.
4. Train on how to politely use the language in various situations: negotiation, problem-solving, dealing with enquiries and answering the phone.
5. Review the grammatical structures of language and writing skill.
6. Develop and administer the test. The test should consist of two parts: writing and speaking with the total score of 100. Each part will contain 50 marks.

The criteria for speaking assessment will consist of listening ability (25 marks) and speaking ability (25 marks). The speaking criteria are:

1. Getting to the point (10 marks)
2. Grammatical accuracy (5 marks)
3. Comprehensible accent and pronunciation (5 marks)
4. Appropriateness of non-verbal language (5 marks)

Referring to the administration and rating procedure of English for Tourism II test at NRRU, the students will traditionally be directly and individually assessed on their speaking ability as required by the course with the course lecturer. It inevitably leads to the problem concerning practicality in test administration which usually takes several days to complete due to a large number of students with the minimum of 230 per course with three lecturers. Another problem is on the reliability of scoring that is

normally rated by the sole rater who is the course lecturer because of the time constraint and number of students.

These practicality and reliability issues remain problematic in several language courses at NRRU. For this reason, the semi-direct web-based speaking test will be developed and administered in this course to entangle the stated problems. Another reason is from the potential of this technological platform regarding its logistic advantage, less reliance on computer sophistication, cost-effective and most importantly in facilitating more realistic test tasks via the use of multimedia presentations (Garcia Laborda, 2007a and 2007b, Roever, 2001 and Hamilton, Klein and Lorie, 2000).

## **2.3 Web-based language testing**

### **2.3.1 The use of web-based in language assessment**

Web-based language testing (WBT) shares similar features as computer-based language testing (CBT) with the evolution of delivery platform, the world wide web. WBT is written in the internet language, HTML and comprises of numerous HTML files. The test takers are required to download these files from the server (testers' computer) to their computer (client) with the use of web browser (e.g. Internet Explorer and Mozilla Firefox) to interpret and display test data (Roever, 2001).

The WBT has more advantages than the CBT in its flexibility and convenience in test administration place and time. WBT is categorized into low-tech and high-tech based on the levels of technical sophistication requirements e.g. the use of adaptive algorithm in test handling and selection, the software, the reliance on computer professionals and the budget (Roever, 2001). The low-tech WBT is the integral part of this study due to the practicality on user-friendly program writing, less dependence on the expert and the budget used in test development.

The WBT is suitable for low-stakes assessment (Chapelle and Douglas, 2006, Roever, 2001), especially for self-assessment e.g. the self-diagnostic standardized test Dialang project due to the limitations on item security and cheating. However, Garcia Laborda (2007b) reports on the use of world wide web platform on standardized high-



stakes tests. He projects that numerous standardized CBTs will eventually be available online and include speaking section. He mentions on the conversion of the two most influential standardized tests, TOEFL iBT and IELTS CBT, that have much influenced on testing context. TOEFL iBT is the online version and globally used whereas the experimental computerized version of IELTS has been used in some countries. The author also presents various test types used in both CBT and online tests ranging from multiple-choice to open-ended productive response. These item types are categorized into four main types and the first one is the test without speaking section: ACT ESL Placement test, WebCape Computer-Adaptive Placement Exam and the Combined English Language Skills Assessment (CELSA). The second category is a computerized test followed by a personal interview. This includes BULATS, an adaptive computerized business test, and the computer-based academic IELTS which is still administered in limited countries as the experimental version. The speaking section is delivered in the traditional face-to-face oral interview. The third type is the use of computer to deliver oral input which includes Basic English Skills Tests (BEST). This standardized test requires the candidates to respond to the oral input delivered via the computer in the speaking section. The last type is the inclusion of the speaking part in the test and TOEFL iBT is in this category. The test assesses all four skills. Double correction system, e-raters and human raters, are used in the writing part whereas speaking is administered via computer and assessed by human. The author concluded that the use of online test technology could be faster, more efficient and cost-effective than the traditional test version; thus, the benefits of this innovative test should outweigh its pitfall. The inclusion of multimedia prompts could facilitate more realistic test tasks for the test takers (Garcia Larboda, 2007a).

Likewise, Hamilton, Klein and Lorie (2000) discuss the feasibility for using WBT for large scale standardized tests due to the numerous technological advantages in inexpensive and rapid scoring, central storage of item banks, and less dependence on sophisticated software and hardware. All these qualifications of WBT are suitable for large scale testing. The authors also mention that the advance of technology made it possible to create new types of questions that can assess complex metacognitive skills. They state a number of issues that are concerned with the effects of this

innovative means on the test takers' proficiency and on groups across course contents. Furthermore, the readiness of infrastructure regarding equity and quality of internet across institutions and the availability of computers should be taken into account. Human resources are another essential issue on the feasibility of using WBT for a large number of candidates that includes both lab staff and teachers. Training and support are important in large scale testing. Another important issue is the budget for test development including score reporting. The authors conclude that this information technology is used in almost all aspects of our life and it will gradually be used for online large scale testing. This includes the use of language learners' communities to facilitate test information and strategies, language test preparation and enjoyment; therefore, further studies are required (Garcia Laborda, 2007b).

Although WBT exploits the advances of technology that can offer interactive test tasks and gear towards ideal learner-centered second language assessment, this test type is still required to undergo validation processes. Chapelle, Jamieson and Hegelheimer (2003) propose validation process for low-stakes ESL online assessment. They stated that test purpose should guide both a validation process and test design. Test Your English, the online low-stakes diagnostic test, was used in the study to investigate the implications of test purpose with 84 EFL/ESL participants. By putting the validation theory into practice, the author exemplified the integration of vocabulary and grammatical acquisition to form the basis of test selection and analysis in the corresponding part of the screen test. They suggested that test impact should be considered as part of test purpose and it could assist in research conceptualization for future studies. They concluded that the investigation on test purpose in validity argument could state some of the factors that affected the experts' judgements on test design and validation which provided the process of putting the theory into practice at the starting phase of test development.

## **2.3.2 Web-based and computer-based speaking tests**

### **2.3.2.1 Web technology for speaking test delivery and administration**

Web technology has been limitedly used in speaking tests (Garcia Laborda, 2007b:8). The example of one of the most prominent online standardized tests is TOEFL iBT which was launched in late 2005. The test was converted into

computerized and online versions respectively. It aims to measure the four skills of non-native speakers of English to use and understand English in academic contexts. Candidates take up to 4 hours to complete the whole test via computer. Due to the objective of the present study, details of the speaking test are emphasized. The TOEFL iBT has been claimed of its improvement to provide a wide range of language tasks that candidates are likely to face in real life setting with the inclusion of integrated tasks (Alderson, 2009). This online version requires candidates to respond to the test through the microphone and their performances are recorded in the speaking part. The test presents the items in a linear way.

The speaking part comprises of six tasks and lasts for 20 minutes. Four of which are integrated tasks that require candidates to respond to what they hear or read whereas the other two are independent and aim to elicit the opinion on familiar topics. Candidates are assessed on their ability to speak about daily familiar topics, summarize, evaluate, compare and synthesize information from multiple sources-the reading and listening passages. In task one and two, candidates hear a statement or a question. They are required to respond to these topics. In task three and four, they first read a short passage and listen to the talk on the same topic. They are required to answer questions from the reading passage and from the talk. In task five and six, they listen to a lecture and are asked a question on the lecture. For the scoring, the rating scale ranges from 0 to 4 and it is based on four criteria: general ability, delivery, language use, and topic and development. General ability involves intelligibility, task fulfillment and coherence, while delivery includes clarity, fluency, pronunciation, intonation and stress. Language use deals with range and control of both grammar and vocabulary. The last criterion, topic and development, deals with the relationship and progression of ideas and relevant content.

Due to the global use of this standardized test over 6,000 institutions and in 136 countries, reliability is important. Alderson (2009) reported the reliability and comparability of the TOEFL iBT based on numerous studies and speaking tests. The statistical analysis demonstrated high reliability at 0.88 by the generalizability theory and 0.84 for the test-retest reliability. However, speaking test tasks contained the evidence of construct-underrepresentation in which only the monologic speaking was

tested. In academic context, candidates were involved in making clear and asking about others' ideas and questioning on what they had heard.

Although the evidence on construct-underrepresentation existed in the speaking part, a 5-year-study of Wall and Horak (2006, 2008a, 2008b cited in Alderson, 2009) reported on the positive effects of the TOEFL iBT that geared towards speaking ability instruction with less evidence on teaching grammar and vocabulary. They found that the impact of the test was less direct while the major factor was from the textbook and instructional methodology. Alderson also praised the development of the TOEFL iBT and pointed out the need for further investigations on the challenging issues of construct-underrepresentation for speaking and writing tasks.

Apart from the being used as a standardized high-stakes test, the WBT is also used for the low-stakes diagnostic purpose and standardized test preparation with a large number of test takers. An example of this online testing is PLEVALEX, a new platform for oral testing in Spanish. The trial model is developed in Spanish and anticipated to be used with other foreign languages. It is developed by Polytechnic University of Valencia, Spain and firstly administered in June 2006 for 500 test takers at the same time and expected to reach 1,500 with the use of technology and mathematic calculation. The test comprises of three sections in which the first two are multiple-choice reading test and composition writing. The last part is speaking that includes three task types: description, semi-direct interview, and answering questions about opinion or information. Test takers are required to attempt the test tasks through the microphone and their responses are recorded and transferred to human raters in the back office. The scores are sent to the central database for grouping and analysis respectively. The trial phase revealed that test takers did not have any problems relating to the test format (Garcia Larboda, 2006).

Garcia Larboda (2007a) reports on the conformity of the PLEVALEX design to the interface design principle proposed by Fulcher (2003b), including test validity and reliability. The author proposes that the test should incorporate interactive audio input that is likely to effectively assess the complexity of language construct.

Additionally, the reliability of the test can be facilitated by the grading system that allows numerous reevaluations. The use of video recording can make it possible for more effective assessment of non-linguistic features than traditional audio recording. Thus, this system has been claimed to enhance the reliability of scoring method in the way that raters could revise, reevaluate and adapt their scoring to optimally conform to common criteria. Consequently, it is expected to reduce both intra- and inter-rater discrepancy. Professors and researchers from 15 educational institutions suggested that the PLEVALEX could provide guidelines to solve the problems in general low-stakes WBT.

Since the WBT for speaking testing is still limited, studies related to the CBT are included in this study to provide a wider range of the integration of technology in speaking assessment.

The Computerized Oral Proficiency Interview (COPI) is the latest version of the speaking test developed by the Center for Applied Linguistics. The previous versions are the Simulated Oral Proficiency Interview (SOPI), the tape-recorded, and the Oral Proficiency Interview (OPI), and the face-to-face interview. COPI employs adaptive technology in test administration with the use of candidates' self-assessment based on the ACTFL Proficiency Level (Novice, Intermediate, Advanced and Superior). The computer selects the four speaking tasks from the bank that are at the similar levels of the self-assessment and the other three are at the higher levels (Jamieson, 2005). The task poses various parts with the integration of written, audio, graphic files and pictures and takes 30 to 50 minutes to complete. Malabonga, Kenyon & Carpenter (2005) investigated on the use of self-assessment, preparation and response time on the computerized oral proficiency test by comparing COPI with SOPI. They found that 98% of the examinees effectively used the self-assessment instrument to choose the task that matched their level of proficiency. However, the comparison of SOPI result indicated that 8% of the examinees selected too difficult task which could be problematic for the starting levels of COPI. Regarding planning and response time, examinees with different levels of proficiency employed different amount of time in which less preparing time with longer responses was found in high proficient examinees. It was also found out that COPI was considered easier than

SOPI, using an adaptive algorithm that could tailor the difficulty level of the test task with an individual examinee (Kenyon, Malabonga and Carpenter, 2001). However, this assertion requires additional investigation, especially on the use of the instrument (Norris, 2001).

The Versant Aviation English Test (VAET) is offered in computerized and telephone version and developed by Pearson Education under the cooperative research with the U.S. Federal Aviation Administration (FAA) to assess candidates' ability to understand and appropriately respond to spoken topic-specific and common English in aviation context under the requirement from the International Civil Aviation Organization (ICAO). Regarding the computer-based test administration, candidates are provided with the microphone, internet connected computer with the Versant Computer-Delivered Test (CDT) software. The test is separated into eight sections with 78 items and lasts between 25 to 30 minutes. Candidates are required to read the printed messages from ICAO phraseology in part one (Aviation English) to the common English in aviation domain in part two. In part three, they listen to the common job-related topics and repeat the sentence aloud. The next part is the short answer question that requires candidates to answer in single word or short phrases to aviation-related topics. This is followed by the readback part in which candidates listen to the radiotelephony messages and say the appropriate readback to confirm their understanding. In part six, candidates hear the radio exchange between the pilot and air traffic controller. They are required to continue the dialogue by correcting information, responding to requests or questions. The following part, storytelling, requires candidates to describe characters, situations and outcomes on aviation situations. The last part, open questions, requires the candidates to answer on aviation-related topics and provide the opinion respectively. This part is not scored. Candidates are assessed by the automatic scoring system that can provide immediate feedback based on the pronunciation, structure, vocabulary, fluency, comprehension and interaction. The test is claimed to pose high reliability ranging from 0.82 to 0.97 for machine scoring and 0.84 to 0.98 for human raters which demonstrated high correlation between the two counterparts. The test also poses high content and

constructs validity (Versant Aviation English Test: Test Description and Validation Summary, 2008:23).

### **2.3.2.2 Web/computer technology for rater training and automatic scoring**

Web/computer technology is not only used for test delivery but also involves in the process of rater training, scoring and speaking test development. In the study of Malone (2007), web-technology intertwines rater training in the test as in Assessment Training Online (ATOL). This online course aims to train professional raters through the more cost-effective alternative means to face-to-face live training. The participants were engaged in the introduction to the ACTFL Proficiency Guidelines Speaking, the core course, with sample rating practices through online reading, chat session, weekly discussion board and electronic assignment submission. After the core course, they were allowed to take the optional language rater training course in Arabic or Spanish. The instructional approach was similar to the core course with the conversion to the target language. They were also required to complete the quizzes. The result from the pilot phase revealed the positive attitudes from the majority of participants regarding effective online distance learning experience with successful written, audio and graphic course materials. They thought that the ATOL course was a cost-effective rater training method that could be used in various locations.

This advanced technology is incorporated in scoring method aiming to provide rapid, consistent and accurate score reports. De Wet, van der Walt and Niesler (2007) investigated the correlation of oral assessment between the Automatic Speech Recognition (ASR), an automated program for large scale oral assessment, and human raters. ASR was implemented with 106 subjects from the Faculty of Education, Stellenbosch University, South Africa in 2006 to provide fast and objective assessment for placement purposes. Students were required to respond to the three task types: reading (read the sentence on the worksheet), repeat (listen and repeat sentences) and open-ended task (unprepared responses to general questions). The test was administered via telephone calls in which students' performances were recorded by a Spoken Dialogue System (SDS) and assessed on their Rate of Speech (ROS). The criteria used for oral assessment were based on fluency, intelligibility and overall

pronunciation quality. Five human raters, ESL or EFL teachers, rated the speeches on the three tasks with an additional overall impression task with a 5-point-Likert scale. The finding revealed that the consistency of human raters was considerably low. The intra-rater correlation values were 0.30, 0.40, 0.72, 0.73 and 0.74. The open-ended task revealed the least correlation among human raters while the highest agreement was in repeat task. In contrast, high correlation for read (0.98), repeat (0.94) and open-ended tasks (0.86) existed in automatic raters. The authors concluded that there was a mid correlation between automatic and human raters ranging from 0.52 (read), 0.58(repeat), 0.48(open-ended) and 0.56(overall impression). These statistical results demonstrated that the automatic rater was likely to rate in a similar way as human raters with high values in the repeat task and low for open-ended task. They pointed out the issue on the use of 5-point-Likert scale that revealed uneven scores from human raters. So, they suggested for finer scale. They also mentioned that the accuracy of the ASR system was at the improving stage by adding the expansion of vocabulary recognition to cope with open-ended task. The ASR was also employed by VERSANT, the LSP standardized test for aviation to provide immediate feedback on score reporting for a vast number of candidates. The acoustic model of the machine was programmed to recognize various kinds of non-native accents. The results of the field test revealed the high reliability of the score reported by the machine (0.93), including the high correlation with the human counterpart (0.94) (Versant Aviation English Test: Test Description and Validation Summary, 2008 pp.23-26).

Moreover, the technology is integrated in test development that included the Computer Assisted Screening Tool (CAST) Framework to investigate and analyze the features of successful oral proficiency tasks. The Computerized Oral Proficiency Instrument (COPI) can provide three times faster rating than the tape-mediated oral assessment (Malone, 2007).

### **2.3.3 Advantages and limitations of web-based language testing**

One of the prominent advantages of WBT is logistic flexibility that allows test administration to happen in any places where the internet access is possible (Hamilton, Klein and Lorie, 2000 and Roever, 2001). Candidates can also take the test



on their preferable time based on the purpose of the test which is mostly used for low-stakes and self-diagnostic that can be seen in the Dialang Project. They also get immediate and specific feedback which is pertinent to some certain aspects of learner-centered second language assessment (Chapelle, Jamieson and Hegelheimer, 2003). Another point that WBT gains the status in assessment context is derived from its user-friendly in test development, especially the low-tech type that does not require expertise in programming and sophisticated hardware and software. The web technology also allows test developers to create the interactive semi-direct speaking test by free software and post the test online on the free platform; hence, the test is considered cost-effective in test production. With the use of the internet system and technology, central storage of test items can be shared and co-developed by numerous test developers within and across the institutions. This global link system enhances collaborative item pool development with the requirement only for the login name and password to access the test. This includes the convenience of test score update and analysis from the central database (Hamilton, Klein and Lorie, 2000 and Roever, 2001).

Although the WBT has been praised of its practicality, this innovative test has many limitations. The nature of the world wide web platform allows everyone with a login name and password to access through the test bank; item confidentiality, thus, is unavoidable and viewed as one of the most influential drawbacks of the WBT. Cheating is another essential difficulty on the use of WBT in which candidates can take the test in an uncontrolled environment. Also, they can ask their peers to take the test for them (Roever, 2001 and Garcia Laborda, 2007b). Technical and electrical limitations include server failures in which candidates cannot get access to the test (Hamilton, Klein & Lorie, 2000 and Roever, 2001). The problem related to adaptivity and compatibility of technical devices and browser can also forbid candidates to download the test due to the different versions of software (Roever, 2001 and Garcia Laborda, 2007b). The issue concerning data storage of the file should be strictly considered, depending on the types and purposes of the WBT. Additionally, technical capabilities e.g. screen capacity and graphic capability, test features and design require additional investigations (Brown, 1997 and Garcia Laborda, 2007b).

## **2.4 The effects of speaking test tasks on speaking performances**

One of the prevalent factors affecting language performances is test tasks (Bachman and Palmer, 1996) which has raised the interest in the testing context including speaking tests. Although this topic has been investigated by a number of research studies (Turner and Upsher, 1995, Lumley and O'Sullivan, 2005, Teng, 2008 and Kim, 2009), the findings remain inconclusive.

Teng (2008) empirically explored the effects of three task types on EFL speaking performance with 30 subjects from Taiwanese universities. The three task types (answering questions, picture description and presentation) were investigated on their effects towards spoken discourse regarding accuracy, complexity and fluency. The author found that there was no difference in performance across the three task types. However, the significant main effects on complexity and fluency were found in different task types focusing on the answering questions task. Subjects achieved higher scores on the answering questions than describing pictures in terms of complexity due to the greater processing demand in task completion. They also did better in answering questions than the other two tasks on fluency due to the greater familiarity within this task type and this structured task posed more possibility to accommodate fluency in speech production. Regarding the subjects' perceptions towards the three tasks, they felt nervous before attempting the task in answering the question parts due to different degrees in communication pressure. The highest mean score was from the picture description task which might be derived from the provision of the visual cue cards; therefore, they stated that this task should be included in all oral tests. Generally, subjects thought that more tasks should be included in this semi-direct speaking test. The author also mentioned that the finding might be affected by the subjectivity in scoring method although the inter-rater was employed in this study.

In addition, task types and context effects via the computerized test on the second language speaking ability were investigated in the study of Kim (2009). The total of 162 adult learners of English as a second language at Teacher College from Columbia University participated in this study. The instruments were the Community English Program test (CEP) in speaking section which was a semi-direct

computerized test and the survey of the test takers' background information. The test takers' performances were investigated on grammatical, discourse, sociolinguistic competence, intelligibility, meaningfulness, and task completion. Multivariate generalizability theory (G-theory) and confirmatory factor analysis (CFA) were used in the data analysis. The findings indicated that the test takers' performances were likely to change according to the context and task types. However, the effects of two factors on the performances were varied. To be precise, the mean differences across tasks were not large. Means of grammatical competence and intelligibility were rather stable across different domains and task types. The small effects of task types were found in some speaking components, sociolinguistic competence and task completion. In addition, the qualitative analysis on the response features across tasks was not notable. The author indicated that the task factors alone insufficiently supported the variation in the test performances.

Similarly, the effects of task types on the relation of Communicative Efficiency (CE) and Grammatical Accuracy (GA) in the direct speaking test were also investigated in a study of Turner and Upshur (1995). CE as referred in the study was the success in communication whereas GA concerned with the accurate use of linguistic forms in the communication events. The two task types (the single sentence creation task and the story retell task) were employed with 130 fifth graders in Montreal, Canada. In the first task, students were required to produce a single sentence from 15 visual cue cards. The second task, on the other hand, required them to watch 2.30 minutes cartoon videotape and retell a story for 1 minute. Participants were rated by four separated rating scales on the two tasks. The finding revealed that the salient difference in the relation of CE and GA was found in the two tasks. For the short utterance task, the relation was linear while the nonlinear existed in the story retell task in which CE exceeded GA. The amount of speech produced was considered the main difference between the two tasks and had the implication on the comprehensibility of participants' speech. In the story retelling task, GA was considered very important. However, in the single sentence production task, high level of CE was related to that of GA; hence, it was likely that GA was necessary for CE in this task. It could, therefore, be concluded that the relation between CE and GA

relied on the task types. In the short utterance task, both CE and GA continually developed, but in the second task CE rapidly rose at the higher level than GA.

Another important factor affecting speaking performance is task difficulty. Fulcher and Marquez-Reiter (2003) explored the effect of task difficulty on EFL speaking performance with 55 students; 23 were Spanish and the others were English speaking students. The authors also proposed the new approach of defining task difficulty by taking into account the first language cultural background apart from task condition, students' proficiency and rater severity. Therefore, the effects of task condition of social power, degree of imposition and first language cultural background on speaking achievement scores were explored in this study. Concerning social authority, each task was marked as high power in the speaker than the hearer, of equal power and less power respectively. The degree of imposition was stated as low and high. The speaking test had six tasks: borrowing books from a professor, asking a subordinate to answer the phone call, requesting for help to move stuff, requesting for a seat on the bus, asking a boss for advance paying, and borrowing a laptop. A pair of students participated in role playing based on the instructions on the card. Each test task took 8 minutes and the students' performances were video recorded and assessed respectively on the successful performance of completing the test tasks. The finding from the univariate analysis of p value revealed the significant effect of both social authority and imposition including the first language background knowledge on score achievement across the two groups' students. The two-way interaction between social power and imposition was significant and the similar result was found in the three-way interaction of the variables. Although the medium effect size was only found in Task one and Task six from the two groups as the indication of the extreme of either social power or imposition, it represented the impact of the first language cultural background knowledge. Despite the finding, the authors argued that change in spoken discourse might not always mean the change of score achievement based on the medium effect size in this study. They stated that ability had greater effects on the change of scores than task conditions in the case that the rating scale was not specific to a certain task. It was suggested that pragmatic features and cultural background should be emphasized on LSP tasks and rating scale design. The medium

effect size of the extreme case in the study was likely to indicate more difficult tasks for certain students which required the specific language background. Thus, L1 background should be included in the construct and this might be suitable for LSP testing. However, this view depended on the purpose of the test. Referring to task difficulty categorization, the authors stated that the use of pragmatic approach in defining task difficulty was sensitive to certain L1 cultural background group; hence, the psycholinguistic approach was more preferable. The authors called for further studies on the varied use of task conditions based on the pragmatic approach.

Similar to the previous studies on the effects of tasks towards speaking performance (Turner and Upshur, 1995 and Teng, 2008), Lumley and O'Sullivan (2005) investigated the effects of task topics on EFL academic speaking performance at the universities in Hong Kong. The finding from this semi-direct speaking test indicated a small effect on the difference in performance across examinees' genders and presenters of the test input (audiences). The total of 894 students who were completing their Bachelor's degrees participated in this study. A Multi-faceted Rasch Analysis was employed to investigate the interaction effects of the task topic, examinees' and audiences' genders. The Graduating Students' Language Proficiency Assessment (GSLPA), an exit examination at the universities in Hong Kong was used in the study. The test was designed to indicate the students' English proficiency to the future employers; the test is therefore considered a LSP testing that it incorporated both social and business English with the integration of speaking and writing skills. There were five tasks ranging from formal to informal situations and consisting of different numbers of examinees. They were provided with prompts indicating the number of audience, purpose of the task and talk, level of familiarity and contextual information. The first task required the examinees to listen to a 5-minute dialogue and make a summary for friends. The second task dealt with advertisement reading and they were required to participate in the interview with the employer. It was followed by the presentation of a graphic task to a group of people. The fourth task required examinees to make a phone message for colleagues and clients and the final task asked them to give advice to the international visitors. Each test task incorporated many skills and their performances were rated on the task fulfillment and relevance,

clearness of the presentation, grammar and vocabulary, and pronunciation on the six levels analytic rating scale. Thirty raters participated in this study. Generally, the analysis indicated a small effect of task topic, gender of examinees and audience on EFL speaking performance. Considering the effect of individual variable, the evidence on the effect of task topic indicated that the bias towards a particular gender on certain topic was found. In Task five, the topic related to horse racing was advantageous in all four criteria for male examinees when the input was presented by male audience. The topic was harder for females due to the interactional effect of the task topic and gender of the audience. However, the topic had the greater effect than the gender of the presenter. It was concluded that examinees were affected to some extent by contextual features of the test tasks; and they also viewed the input as the communicative event with the appearance of interaction, even when there was no interlocutors during the test administration. The authors referred to task difficulty that the finding in this study was pertinent to the previous research findings, especially to the investigation of Fulcher and Marquez-Reiter (2003) in which the task might affect individuals differently.

Despite the potential effect of test tasks on the performance and the tremendous impact on testing context, the standardized academic TOEFL prototype speaking and writing tasks have been undergone the field-testing phase to ensure the positive effects on test takers' performance. This exploratory research was conducted by Cumming et al. (2004) to evaluate the test content validity, correspondence of test performance to that of actual class performance and its relatedness to the educational context in North America setting. Seven professional EFL instructors from three universities were selected and they were required to rate their students' abilities in the actual classes and the TOEFL test performance on the 5-point scale. They were also required to evaluate the overall effectiveness of the prototype tasks on the questionnaire and participation in the interview. Seven tasks were piloted in which three were writing and the others were speaking. The speaking tasks comprised of independent speaking tasks that asked examinees to speak about their personal experience, and the rest involved academic topics in which they responded to listening (lectures and conversations) and writing (reading passages) stimuli. It was

found that all the instructors praised the prototype tasks “to be realistic and appropriate in their simulation of academic contents and situations, in the skills they required the students to perform, and in the opportunities they provided for students to demonstrate their abilities in English” (Cumming et al., 2004:134). Essentially, the correspondence between the test task performance and the actual ESL class was found in this study as indicated by the significant correlation in the independent speaking task,  $\rho=.48$  ( $p<.01$ ). Although the teachers had positive impression on both independent and integrated speaking tasks, criticism had been made on the latter task type. The most outstanding problem was derived from tasks reliance that made them difficult for less proficient students who could partly understand ideas, vocabulary and contexts in both written and listening prompts. The instructors suggested that the tasks could be better developed by selecting topics, prompts and task types that were familiar to all students and the visual and graphic inputs should be provided to accommodate the students. Some of the instructors thought that the tasks could be more challenging by requiring students to use their cognitive ability to attempt the test. They also suggested on the verification of the reading-speaking task that a fundamental level of rhetorical and schematic organization of the information on texts should be provided. The authors stated on the limitation of the study on the number of participants. Concerning the claim for construct validation, the evidence could not be used due to the insufficient data from both teachers and students’ performances.

### **2.5 Test takers’ attitudes towards web-based/computer-based speaking tests**

As part of test takers’ characteristics that have potential effect on the language performance (Bachman and Palmer, 1996), test takers’ attitudes, particularly on the innovative mode of test format are important in the testing context. Despite the similar feature of the test format which is referred as the previous version of the web-based speaking test (WBST) and the limitations of research studies investigating test takers’ perceptions on the internet-based format, studies on the computer-based speaking test (CBST) are reviewed. Some studies investigated test takers’ perceptions towards CBST in comparative studies of the computer-based format with the precursor version of tape-based tests and that of traditional face-to-face tests (Kenyon and Malabonga, 2001, Norris, 2001, Kenyon, Malabonga and Carpenter, 2001 and Warschauer, 1996).

The findings from these studies indicated test takers' preference on the computerized version to that of tape-based and traditional face-to-face speaking tests.

A comparative study on test takers' attitudes towards the CBT and other oral proficiency assessments in three languages, Spanish, Arabic and Chinese, was conducted by Kenyon and Malabonga (2001). The tape-based Simulated Oral Proficiency Interview (SOPI) and Computerized Oral Proficiency Instrument (COPI) developed by the Center of Applied Linguistics were used with 55 graduates and undergraduates and 24 Spanish participants were required to attempt the Oral Proficiency Interview (OPI). After completing the test, the questionnaires were administered to participants to investigate their attitudes on the tests in six aspects: adequacy of the test in reflecting strengths and weaknesses of language proficiency, difficulty, fairness of speaking situations, test taking nervousness, clarity of directions and accuracy of the test that could be generalized into real life setting. The results indicated that participants viewed the COPI to be more advantageous than the SOPI in the application of adaptive algorithm that could match the difficulty level of tasks with their proficiency level. Therefore, the computerized version was considered easier than the tape-based version, particularly for the low proficiency group. They also thought that the innovative COPI was fairer and less nervous than the SOPI. Concerning the three test forms, Spanish students felt that both SOPI and COPI were similar to the OPI with the exception that the original version was more advantageous in assessing real life speaking skill in which these two semi-direct tests could not replicate the interactive face-to-face interview. It was concluded that the use of the test formats depended on the criteria use for evaluating the performance. The findings showed that the technology-based tests could capture features of speech production without the interaction nature of the conversation. Therefore, the innovative tests might not be applicable in the context where interaction features were important in the test.

Referring to the instrument used in this study, Norris (2001) questioned on the validity of the questionnaire that did not provide the examinees with adequate options on test format's appropriateness e.g. the option as "neither of tests was appropriate". Therefore, the finding was unclear on the number of examinees who would choose



this option. He also pointed out that although the research design in the present study was inappropriate; the findings showed that examinees felt equally comfortable to the computerized version. He praised the advantage of the COPI concerning rating process facilitation and examinees' performances elicitation with the advent of technology. However, he cautioned that computerized context features might have negative effects on the examinees' perceptions on language tests; consequently, this might change their performance. He requested the additional studies for the COPI and this was agreed with Kenyon, Malabonga and Carpenter (2001). The authors supported Norris's view that with or without the use of computer or technological support, speaking assessment in second language was complex and the evidence-centered design was critical.

Similarly, the findings from Warschauer (1996) revealed subjects' preference for the computer-based discussion task to that of face-to-face interaction in the second language classroom. Sixteen subjects participated in the study and they were required to report on their preference for the particular mode of task delivery. Results from the survey indicated more positive attitudes on the electronic communication than that of the face-to-face version. Participants felt free to express themselves and they also reported to be more comfortable and creative in sharing their opinions. They were at ease and thought that the computer-based discussion facilitated their thinking ability and the program was user-friendly. In terms of language complexity, quantitative analyses on lexicon and syntax were employed. The findings showed that the semi-direct discussion tasks were advantageous in eliciting the complex language structures which was pertinent to Norris's view (2001). Furthermore, longer turn-taking was found with less direct interaction in the electronic discussion while a number of confirmation checks and active responses were found in the face-to-face version. On the aspect of language formality, more informal expressions were found in the face-to-face mode whereas the more formal use of language appeared on the electronic mode. It was concluded that despite the benefits from more complex language production in the electronic mode, it may accommodate learners to gain more formal and complex communication skills. The author stated on the limitations of the study on the use of scale assessment and thus called for future studies.

## **2.6 Strategies used in speaking tests**

### **2.6.1 Taxonomy of strategies used in speaking tests**

Until present the consensus on strategy taxonomy has not been settled as claimed by a number of studies (Cohen, 2006 and Swain et al., 2009). This issue remains as one of the challenges in test taking strategies field (Cohen, 2006). In language testing studies, the main emphasis of strategy is on the test taking strategies used by learners to perform the test tasks involving with the communication need during test taking process (Swain et al., 2009).

Test taking strategies is defined by Cohen (1998), one of the most influential scholars in test taking strategies that test taking strategies are made up of language use strategies and test-wiseness strategies. Language use strategies deal with selective and conscious use of retrieval strategies (to retrieve information for attempting the test), rehearsal strategies (to rehearse before giving the actual information), cover strategies (to make the answer look good) and communication strategies (Cohen, 1998:219).

On the other hand, test-wiseness strategies rely on test takers' knowledge on the way to achieve the test. These strategies involve matching similar information, matching across items, and taking shortcut to answer the test task. There is a tendency that the high achievement score on the test is derived from the effective use of these test taking strategies not the actual language ability. However, the frequent use of the strategies cannot be considered as the indication of task achievement. Effective use of certain test taking strategy depends on the way the test takers use a particular strategy in a particular time and in a certain task with their "particular cognitive style profile and degree of cognitive flexibility, their language knowledge and their repertoire of test-taking strategies" (Cohen, 1998:220).

Following the interactional view of language competence model (Bachman and Palmer, 1996), strategy used in speaking is derived from analysis of strategic competence that "provides details of what we could look for in test-takers' speech."

(Fulcher, 2003a:31). Strategy used in speaking comprises of two main categories: achievement and avoidance.

Achievement strategies are used when learners lack knowledge in the target language, particularly lexicon and grammar. Therefore, they try to overcome this difficulty by finding way around the problem. This category includes six strategies: overgeneralization/morphological creativity, approximation, paraphrase, word coinage, restructuring, and code switching. Overgeneralization/morphological creativity strategy is used when learners do not master in the target language; so they are likely to transfer a particular target language system to all structures e.g. the use of past tense ‘-ed’ to all verbs. Approximation strategy is used in the case that learners replace more general words to unknown words. Paraphrase is the use of near synonym words for the words they need. This includes the use of circumlocution. Word coinage relates to the invention of new words for the unknown words. Restructuring relates to the use of different words or grammatical structures to resend the message that is not comprehensible. Learners use cooperative strategies to get help from the listener or interlocutor, particularly in saying the unknown word. Code switching is the use of shared language to overcome the lack of the target language knowledge. Non-linguistic strategies deal with the use of gestures to elicit language or assist in communication. The last three strategies are used only in the face-to-face communication and thus considered social interaction strategies.

Avoidance (or reduction) strategies, on the other hand, are used when learners do not have full control of the target language. They employ these strategies when they have linguistic means to convey. This category of strategy has been used mostly by low proficiency level learners (Littlemore, 2003). Avoidance strategies are made up of formal and functional avoidance. Formal avoidance occurs when learners avoid using a particular part of the language system. However, this strategy is hard to detect. An example that deals with this strategy is the avoidance of using passive voice. Learners used functional avoidance to avoid the particular semantic, topic and end the conversation. The latter two are quite serious compared to the first type. Semantic avoidance occurs when learners do not know particular lexicons in the target language.

For the present study, strategy used in the speaking test is derived from reported behaviors or thoughts during the test taking process. Theoretically, test taking strategies are “the conscious, goal-oriented thoughts and behaviors test takers use to regulate cognitive processes, with the goal of improving their language use or test performance” (Swain et al., 2009:2). Therefore, test taking strategies in this study incorporate mixed frameworks and comprise of communication strategies (achievement and avoidance strategies), cognitive strategies (selecting, comprehending, storing memory and retrieval) and metacognitive strategies (goal setting, organizing, planning and evaluating) (Fulcher, 2003a, Cohen, 1998, Bachman& Palmer, 1996 and Swain et al., 2009).

### **2.6.2 Studies on strategies used in speaking and oral tests**

Research on strategies used in speaking and oral testing was conducted on the relationship between strategies used and language performance with the varied results. Although participants employed a wide array of strategies, only some strategies had positive effects on language performance (Song, 2005, Lam, W.Y.K., 2007). This includes the interaction between strategy used and level of proficiency; however, the findings remain unclear (Nakatami, 2006, Hong-Nam & Leavell, 2006, Ting & Phan, 2008, Cabaysa & Baetiong, 2010, Ma, 2009 and Mendez Lopez, 2011). Some studies emphasize the effectiveness of certain strategies used for particular participants with varied findings (Littlemore, 2003 and Wiggleworth, 1997 and Anderson, 1983 cited in Van Moere, 2006).

A study conducted by Song (2005) explored the language strategy used in the Michigan English Language Assessment Battery (MELAB) and the relationship between the strategy used and learners’ test performance. The total of 179 test takers who took MELAB participated in this experiment. The test aimed to assess the English proficiency of non-native adult learners. The test comprised of three parts: composition, listening and written test with the option of speaking. The speaking section was a direct test that required test takers to speak to local examiners for 10 to 15 minutes. They were assessed on fluency and intelligibility, grammar and vocabulary; interactional skills included functional language use or sociolinguistic

proficiency. Questionnaires on strategy used focusing on cognitive and metacognitive strategies were distributed to gain the data on reported strategies. The findings revealed that the test takers used six types of cognitive strategies: repeating/confirming information strategies, writing, practicing, generating, applying rules and linking with prior knowledge. They also used three metacognitive strategies: evaluating, monitoring and assessing. The analysis indicated that the more test takers synthesized what they had learned and applied it in the actual practice, they performed better. In contrast, the use of repetition strategies led to worse performance. Regarding one of the cognitive strategies, generating, it was found that it had a significant negative effect on writing, listening and the total score, but it did not have any effects on Grammar, Vocabulary, Cloze and Reading (GVCR). It was hard to demonstrate any interpretations that monitoring solely predicted GVCR scores. Certain strategies such as applying rules, practicing, assessing and evaluating did not have any impact on the performance. The results of this study were pertinent to those of previous investigations and it was concluded that language performance was not enhanced by every strategy used. Some strategies had significant positive effects while the contrast was found in some particular strategies. Some strategies had no effect at all on language performance. Although the finding indicated a linear relationship between strategy used and language test performance, the effect was weak. Apart from language strategy used, other factors, e.g. communicative language ability, characteristics of test takers and test methods or test tasks proposed by Bachman (1990) could affect language performance. Reported strategy was only part of test takers' characteristics. Additionally, the limitations of this study were stated on the nature of strategy used, methodology and the use of only self-reported questionnaires that might affect validation of the study. More comprehensive investigations on strategies used and their relationship with language performance are recommended.

Similar to Song (2005), Lam, W.Y.K. (2007) conducted an in-depth study on ESL learners' strategies used and problems in oral communication. Eight ESL learners in which four were from the experimental group and the rest were from the control group participated in the 10 minutes group discussion task. At the end of the

session, they were interviewed with the use of the stimulated recall approach. They were required to watch the videotape recorded during the discussion and required to report on their thoughts while they were performing the task. The finding showed that the participants employed the total 19 individual strategies in which from two to seven types were reported by each participant. The most frequently used strategies were paraphrasing (15 times), simplification (ten times) and activating background knowledge (eight times). Two participants who were the most and least articulate were selected to compare the strategies used. An analysis of the stimulated recall showed that they used different strategies to resolve the problem during the task in which a high articulating participant could effectively employ metacognitive strategies in planning, monitoring and evaluating the task. Due to the small number of selected participants, the limitations of this study should be considered. It was concluded that knowing the learners' problems and strategy used in oral communication helped both teachers and learners to understand and develop the oral skill. Stimulated recall was a potential methodology to provide insightful information. Concerning the validity issue, the use of observation data could be employed to strengthen the findings derived from such method.

In addition, learners' strategies used in communication tasks were also investigated by Nakatani (2006) as part of the validation study of the oral Communication Strategy Inventory questionnaire (OCSI). The comparison of strategy used across proficiency levels was conducted and the finding showed that some strategies were favored by a particular proficiency group. From the first phase of the OCSI development, eight speaking categories with 32 items were selected and included. They were social affective, fluency-oriented, negotiation for meaning while speaking, accuracy-oriented, message reduction and alteration, non-verbal strategies while speaking, message abandonment and attempt to think in English. Sixty-two Japanese female students were selected to participate in the study and were separated into high, middle and low proficiency groups from the Oral Test Assessment Scale for Japanese ESL. Only high and low proficiency groups were requested to attempt a simulated conversation task. Upon the completion of the test tasks, the participants were required to report on their strategic behaviors on the OCSI. They were also

asked to report on the Strategy Inventory for Language Learning (SILL), a self-report questionnaire invented by Oxford (1996, cited in Nakatani, 2006). It was found that social-affective strategies, fluency-oriented strategies and negotiation for meaning were used more by the high proficiency group than the low proficiency one. Among the three strategies, negotiation for meaning strategies posed the highest difference across proficiency levels. Additional investigations on this issue were required. High proficient students reported using social-affective strategies for affective factors control and they also maintained the flow of conversations by using fluency-oriented strategies. The author suggested that these strategies should be included in the strategy training as they were found to be effectively used by proficient EFL speakers. However, an in-depth investigation was needed due to the limitations of the present study on other variables that might affect reported strategy e.g. gender of participants. Moreover, additional data collection methods should be included to compensate for the validity issue of the questionnaire and the methodology used in the present study. Also, the studies in other foreign language context should be investigated before drawing on the generalization of the results.

Likewise, Hong-Nam & Leavell (2006) explored the language learning strategy used of ESL students in integrated skills Intensive English Program (IEP) focusing on the relationship between language strategy used and the target language proficiency, and assessed the strategy used across genders and nationalities. The total of 55 ESL students participated in this study and they were classified into three groups: beginning (11), intermediate (30) and advanced (14). They were diverse in linguistic backgrounds. Japanese was the largest group (40%), followed by Taiwanese (22%), Korean (20%) and others (18%). They were asked to report on language learning strategy used in four skills through the Strategy Inventory for Language Learning (SILL). A self-report questionnaire emphasized on six categories: memory, affective, cognitive, compensation, metacognitive and social strategies. They were also required to fill in an Individual Background Questionnaire (IBQ). Generally, it was found that memory and affective posed a statistical difference when compared to the other four categories ( $M= 3.4$ ). In other words, these two strategies were used least by the participants. Metacognitive ( $M=3.66$ ), social ( $M=3.62$ ), compensation

( $M=3.59$ ) and cognitive ( $M=3.44$ ) were favored by the participants in respective orders. It was discussed in the article that affective strategies were least reported due to the fear of making mistakes. It was possible that the IEP training might have an impact on the preference of strategy. This included the infrequent use of memory strategies due to the difference in definitions in the study that dealt with presenting new vocabulary, employing rhymes and creating a mental or spatial image which was not related to rote memorization. Regarding the language learning strategies used and level of target language proficiency, the results revealed that the intermediate learners employed more strategies than their counterparts. The statistical difference was in the use of compensation strategies. These strategies were favored more by the middle level learners than the advanced group. Contrastively, metacognitive strategies were used by the beginning and intermediate groups whereas the advanced level preferred social strategies most. Affective strategies were used least by the beginning and intermediate groups. Memory strategies were also used least by the advanced and beginning groups. Due to the contradict findings of this study to the previous research, it could be explained that the intermediate learners gained sufficient knowledge and were ready to move beyond their actual performance and they were conscious on what they were learning and this contributed to the more report on strategies used. For the beginners, they have not reached sufficient knowledge on their target language learning and how to apply effective learning strategies. As for the advanced level, their learning process became automatic and intrinsic; therefore, they reported the lower use of strategies. Regarding the strategy used by gender, there was no statistical significance across genders, but a statistically significant difference existed in the use of affective strategies which were used more frequently by females. Male participants preferred metacognitive and compensation strategies most and affective strategies least. Metacognitive strategies were also the most preferred strategies used by females whereas memory was the least. Additionally, there was a statistically significant difference in the use of metacognitive strategies between Japanese and others. Metacognitive strategies were used most by Japanese, Korean and others whereas social strategies were most preferred by Chinese. Memory and affective strategies were used least by Chinese while affective strategies were least used by Japanese and others. Memory strategies were least reported by Koreans.



Moreover, Cabaysa & Baetiong (2010) compared the frequency and types of language learning strategies of intermediate and novice students. The authors found that the two ability groups reported using metacognitive, social/affective and compensation strategies respectively. There was a significant difference in the frequency of the metacognitive strategies between the two groups (Novice=75, Intermediate=165). The intermediate group reported more than twice as many metacognitive strategies as the novice group. Within metacognitive strategies, planning was used most because this sub-strategy was an essential part of the process of the speech production. It could enable the subjects to intentionally or consciously focus attention on a task, engage in planning and monitor their output, and also lead them to assess their performances.

In addition, the correlation between proficiency levels and the types of strategies were found in some studies (Ma, 2009 and Mendez Lopez, 2011). Ma (2009) conducted an empirical study on the comparison between strategies in oral English classes and writing classes. Cognitive, metacognitive, social and affective strategies were investigated. As part of the study, it was found that proficiency levels were significantly and closely correlated with the use of speaking strategies. Students with high proficiency levels tended to employ more language strategies than those with low proficiency levels. Similarly, these findings correspond with the recent findings of Mendez Lopez (2011). The authors investigated on the speaking strategies used by BA ELT students from five public universities in Mexico which were categorized into beginner, intermediate and advanced groups. The strategies included asking for repetition, paraphrasing and asking for message clarification. It was found that asking for message clarification was the most reported strategies from the three proficiency groups. More specifically, students selected strategies in relation to their proficiency levels. Beginners used speaking strategies to compensate for the lack of language knowledge whereas intermediate students were more likely to make use of their background knowledge; and thus, this group reported the most varied strategies among the three groups.

Contrastively, Ting & Phan (2008) found no difference in the total number of strategies used in the oral simulated discussion task across EFL learners' proficiency

levels. The authors investigated on the influence of the target language proficiency of EFL learners and interlocutors on the use of communication strategy with 20 undergraduate Malaysian participants. They were separated into proficient and less proficient groups by the Malaysian University English Test (MUET) scores. They were grouped into pairs and required to participate in a simulated oral discussion task. Their performances were transcribed to identify the frequency of strategies used in 13 types of strategies. The finding showed that the proficient group employed 132 individual strategies whereas the less proficient group used 142 strategies. Thus, the statistical data revealed no difference in the total number of strategies used. In general, the participants relied mostly on L2 strategies with the use of 92 times out of the total 274. However, less proficiency group depended mostly on L1 strategies, especially on language switch. Restructuring (85 times), lexical repetition (61 times) and tonicity (46 times) were most frequently used strategies by the two groups. However, a slight difference in frequency existed in the restructuring strategy which was used by the proficient group 44 times whereas the less proficient group used 41 times. In lexical repetition strategy, the less proficient group used this strategy 33 times for clarification, emphasis and topic maintenance while 28 times were donated for their counterpart in topic maintenance, topic salience marker and clarification. Furthermore, the finding indicated that the proficient group was potentially influenced by the use of tonicity and language switch. The frequency showed that proficient group employed this strategy 36 times whereas only 10 times were used by the less proficient group. Tonicity in this study dealt with the ability to use stresses and pitches to point the key information and mark the given and new information. This strategy indicated the proficiency in the target language. Contrastively, language switch strategy was used most by less proficient participants for 25 times while the high level participants used this strategy twice. The frequent use of lexical repetition by the high proficient participants to continue the dialogue was found in the conversation adjustment.

Moreover, Littlemore (2003) investigated the effectiveness of compensation strategies which is one type of communication strategies used by synoptic and ectenic learners. Compensation strategies were used when learners lacked linguistic

competence in the target language. From previous findings, people's cognitive styles influenced their communication preference. Synoptic cognitive learners relied on their preconscious and preferred comparison-based strategies while ectenic cognitive learners needed conscious control of what they were learning and emphasized on the individual features of the target language. The total of 82 French speaking students of English at the university level participated in the study and their English proficiency level was approximately at band three from the Interagency Language Roundtable (ILR). Participants were asked to attempt the communication task with 15 written words in French and were asked to explain on what they saw in English. Their performances were recorded and listened by English speaking raters. All recordings were transcribed and analyzed on the frequency of strategies used based on Poulisse's (1993 cited in Littlemore, 2003) strategy taxonomy which included substitution (six strategies), substitution plus (one strategy), reconceptualization (five strategies) and functional reduction (two strategies). Criteria for judging the effectiveness of strategies were on ease of comprehension, stylishness of expression and proficiency. Correlation was calculated between the score of communication effectiveness and frequency of strategies used. The finding showed that strategies that were preferred by ectenic learners were more communicatively effective than synoptic learners. The most effective strategy in this study was reconceptualization. It was followed by substitution, substitution plus and functional reduction respectively. It was suggested that ectenic learners tended to be more benefited than their counterparts in the situation that required the communication strategies to compensate for gaps in the target language knowledge, particularly when the context was not shared or understood by the interlocutors. For the ectenic learners, shared knowledge and imagination were required. The findings also indicated that functional reduction had negative correlation with all three criteria of communicative effectiveness while the reverse was true for reconceptualization. The most effective strategy in this positive strategy was componential analysis. Regarding individual strategy in substitution, conventional analogical/metaphoric comparison and literal comparison were the most effective strategies that enhanced the language to be clearer, more stylish and more proficient. Within substitution plus category, morphological creativity significantly correlated with the proficiency level in the way that this strategy made the learners

more proficient. One limitation of the present study was on the finding of the preferred strategy by ectenic learners that might not be universal. This included the generalization of the findings that might not be applicable to all language learners due to the use of test items, the narrow focus on compensation strategies and the use of transcripts during scoring method that affected the naturalness of communication.

Additionally, as part of the metacognitive strategies used in speech production, the effect of planning time on the oral achievement regarding complexity, fluency and accuracy between the high and low proficient examinees was investigated by Wigglesworth (1997). Although there was no statistically significant difference of the planning time's effect on oral performance, the author found that the proficient examinees benefited to some extent. The total of 107 subjects were randomly assigned into groups A and B and categorized into high and low proficiency groups based on the scores derived from the tape-based Australian Assessment of Communicative English Skill. The subjects were asked to attempt eight tasks with 1 minute planning time per task and the scores from each task were compared concerning with and without planning time. The results showed that the high proficiency group benefited from the planning time only in the difficult tasks that required cognitive competence whereas in the basic tasks such requirement was unnecessary. Therefore, cognitive load played a significant role in attempting the test tasks. On the other hand, the low proficiency group might not use the planning time effectively and it was likely that they made use of the planning time focusing on content not on the grammatical accuracy of their speech. On the aspect of complexity of speech production, discourse analysis on the percentage use of subordinate clauses revealed that planning time was advantageous for the high proficient group only in the difficult tasks while no benefit was found for the low proficiency group. Regarding fluency that involved less frequency use of self-repairs and repetitions, the results partially indicated significant benefits for the low level group on the test. In terms of accuracy, the investigation was on the use of plural '-s', verb morphology and articles, the finding showed the tentative effect of planning time on some measures in the high proficiency group particularly on the most difficult task. In contrast, the low proficiency group benefited from planning time on the use of indefinite articles.

Based on the findings, it was concluded that planning time might be appropriate only on more difficult tasks and it was advantageous for high proficiency examinees. These findings agreed with the result from the study of cognitive development from Anderson (1983, cited in Van Moere, 2006) in that a particular proficiency group was benefited from a particular strategy. The author indicated that the mastery in the target language linguistic knowledge facilitated the candidates' cognitive processing resources to be able to effectively produce the speeches. For this reason, the high proficiency candidates may be advantageous from their cognitive processing as they had to spend less time on the input processing and responses formulation than the low ability candidates.

### **2.6.3 Studies on strategies used in web-based/computer-based speaking tests**

A number of research studies on technology-integrated speaking tests showed that examinees employed a wide range of strategies used to attempt the test tasks (Swain et al., 2009, Smith, 2003 and Shohamy, 1994). Some studies focus on communication strategies used (Smith, 2003 and Shohamy, 1994) whereas some include the effects of tasks on strategies used (Swain et al., 2009 and Smith, 2003). The findings regarding the effects of task types on test takers' choices of strategies were unclear because some task effects were found in some particular studies (Swain et al., 2009).

Swain et al. (2009) investigated on the strategies used in the internet-based TOEFL focusing on speaking section across the three task types: 1) independent, 2) integrated reading, listening and speaking and 3) integrated listening and speaking. The relationship between reported strategies, test takers' performance on the task based on their test scores was examined. The participants were 16 graduates and 14 undergraduate Chinese speakers. They were separated into four proficiency levels from the language proficiency pre-test scores: graduate advanced (A), graduate intermediate (B), undergraduate advanced (C) and undergraduate intermediate (D). The background questionnaire was distributed to all participants before the administration of the Speaking Section of the TOEFL iBT (SSTiBT). The test

comprised of six tasks and was classified into three task types from the skill requirement. Task group A was the independent speaking task, followed by Task group B integrated reading, listening and speaking and finally Task group C integrated listening and speaking. Participants were asked to do the familiarized version of SSTiBT which used the actual administration time 1 week prior to the research version. Task C was used in the study to elicit reported strategies used in English or Chinese for the stimulated recall with pauses at the end of each task. The findings showed that participants used overall 49 individual strategies across six tasks which were categorized into five types: metacognitive, communication, cognitive, approach and affective. Metacognitive (33.42%), communication (26.48%) and cognitive (25.04%) were reported as the three most frequently used strategies. However, the relationships among the strategy categories (communication and cognitive, approach and metacognitive and communication and metacognitive) were negative. In other words, participants who used more communication were likely to use less cognitive strategies which were similar to the rest two cases. The ten most frequently used sub-strategies were cognitive: mechanical means to organize information (11.68%) and attending (5.17%), communication: organizing thought (7.46%) and linking to prior experiences/knowledge (6.06%), and metacognitive: planning (5.88%), evaluating the content of what was read/heard (5.45%), monitoring (5.19%), evaluating performance (4.23%), setting goal (3.95%) and evaluating language production (3.78%). Six strategies were in metacognitive whereas the other four were in communication and cognitive. Regarding strategies used and proficiency levels, communication strategies were significantly used more by undergraduates, particularly in the organizing thought strategy. Contrastively, graduates significantly employed more cognitive (attending strategy) and affective (justifying performance strategy). The results of the study indicated no statistically significant differences in the strategies used across the proficiency groups. Furthermore, participants employed similar strategies within the integrated tasks (Group B and C) than the independent task. They also used a wider range of strategies in the integrated task than in the independent one. The only significant difference was found in the use of approach strategies between Task one and Task two which was in the Task group A. It was concluded that although the relationship between strategies used and test performance

depended on the complex interactions of participants' characteristics, tasks and context; reported strategy was important and should be considered as part of communication performance construct. The authors stated on the limitations of this study on the aspect of real examination condition and the list of strategy was the partial representation of possible reported strategies used. Also, the small number of participants could lead to no significant difference in the finding. Another point was derived from the unsettled taxonomy of strategy which was considered the general weakness on studies of strategy. Therefore, additional studies were required.

Likewise, the strategies used in computer-mediated context were investigated by the Smith (2003). The focus of the study was on communication strategies for maintaining discourse or conversation as well as compensation strategies which were used to compensate for the target language knowledge. The effects of task types on strategies used and the effectiveness of particular communication strategies were also explored. Participants were 18 intermediate ESL in the U.S. universities with diverse backgrounds in L1. They were required to participate in the pair task in the computer-mediated communication environment (CMC) called ChatNet, a software program used for message typing communication for 5 weeks. From previous studies, CMC shared similar characteristics as spoken communication and the benefit gained from this technology-integrated communication was likely to transfer over to spoken language (Smith, 2003). A week prior to the experiment, participants were pre-tested and during the experiment; they were asked to attempt the jigsaw and decision making tasks through online chats in dyads within 30 minutes for each task. Their chat scripts were transcribed and the post-test was administered. It was found that the participants used a wide range of communication strategies during the CMC task performance. The four most frequently used out of 26 strategies were: fillers (54 times), substitutions (43 times), framing (30 times) and politeness (28 times). Fillers were used for conversation maintenance as pauses and time gaining strategies as in face-to-face conversation. Framing was used as explicit marks of topics and turns in written discourse and as intonation and pitches in oral discourse. The finding indicated that the CMC environment required learners to explicitly signal the transition clearer than that of face-to-face communication. Substitution strategies were

used since the learners were required to use the text-based surrogate with the interaction of their low level proficiency as well as the written context of the CMC that encouraged learners to use abbreviations e.g. I C for I see. Due to the CMC environment, the learners were opted to be explicitly polite to ensure that their partners engaged in cooperative behaviors. The present study found no task effects on communication strategies; however, the task types might affect the use of compensation strategies. Almost twice as many compensation strategies were used in decision making task (65%) compared to those of jigsaw task (35%). The data showed that single strategies outnumbered the multiple strategies. Interestingly, the learners did not use L1 transfer strategies at all in the decision making task whereas they used it once in the jigsaw task. It is worth noting that the finding showed that compensation strategies used in the present study were equally effective on the learners' post-test performance which was contradictory to the face-to-face studies. In the latter environment, mixed strategies were the most effective and the single conceptual-holistic strategies were the least effective. Although there was no statistical significance in the use of orientation strategies, it should be noted that they were more frequently reported than other strategies and were almost entirely found in the jigsaw task. Further investigations on the role of more task types in communication strategies were recommended.

With the similar features as semi-direct oral tests, the communication strategies used in the SOPI was investigated with the direct version of OPI by Shohamy (1994) as part of the validation study of the two tests. Empirical evidence was derived from 20 transcripts analyses in which ten were from OPI and the rest were from the semi-direct version of SOPI. Frequency of strategies used across the two test forms was used in the study. Oral strategies in the study were shift of topics, hesitation, self-correction, paraphrasing and switch to L1. The significant difference was found in paraphrasing. On SOPI, examinees employed self-correction and paraphrasing more frequently than switching to L1 while the last strategy was used more in OPI. It was stated in the summary that self-correction strategy was used more frequently in SOPI which represented more attention on linguistic accuracy to monitor their language. The contrast was found in SOPI with more focus on how



language was produced with more frequent use of self-correction, especially on grammatical accuracy. The distinctive frequency of paraphrasing in the semi-direct test was derived from the specific question in which respondents made use of paraphrasing when they could not use the similar lexicons. In contrast, a question on OPI was less specific and respondents could also find ways to answer or even shift the topic. Furthermore, the lack of immediate feedback from the interlocutors was another reason that respondents made use of paraphrasing to insure the transmission of the answers. On the other hand, more frequent use of switching to L1 strategy in the direct oral test indicated insurability that the interlocutors understood the message. It was found that this strategy occurred in the explanation and mostly in the single word. Another explanation was from the occurrence of interlocutors. The respondents adapted their speeches by using a number of strategies, particularly the switching to L1 strategy. In contrast, the SOPI respondents were more concerned with completing the tasks than adapting their speeches since there was no interlocutor involved during the test administration.

## **2.7 Studies on research methodology**

### **2.7.1 Stimulated recall methodology**

Stimulated recall (SR) methodology is the sub-category of introspection methodology that is widely used in second language research to investigate learners' thinking process and strategies. Learners are asked to report their thoughts while they are performing the task or activity after they have completed them. The prominent feature of SR is the use of a stimulus to support the recall (Gass & Mackey, 2000, 2007, Lye, 2003, Edwards-Leis, 2007 and Polio, Gass & Chapin, 2006). The example is the use of videotape to elicit learners' thoughts while they are carrying out the task which is employed in a number of studies (Gass & Mackey, 2007, Lye, 2003, Edwards-Leis, 2007, Polio, Gass & Chapin, 2006 and Swain et al., 2009). Verbal report is included in SR methodology (Gass & Mackey, 2000, 2007). The potential advantage of SR is in the utilization of naturalistic context, specificity of the domain and the holistic approach to the cognitive process (Lyle, 2003:871). However, as similar to other methodologies, SR is not without the limitations, particularly on the

problems that occur from memory and retrieval, timing and instruction. It is still regarded as a valuable tool that should be used with the well-planned recall designs (Gass & Mackey, 2007). Therefore, it is recommended that the data should be obtained as soon as possible after the activity is carried out to increase the possibility that the data are from recent memory. Another point is on the use of a stimulus that should be effective in activating the memory. During the data collection procedure, the intervention by the researcher/interviewer should be kept to minimal. In other words, the participants should not be cued during the procedure. Pilot study is useful in this case. There should also be a strong link between the amount of structures and research questions to gain the useful data (Gass & Mackey, 2007, pp.54-55 and Lyle, 2003). Additionally, Lyle (2003) substantively gathered the best practice techniques of SR and stated that anxiety should be reduced during the procedure and the perception of judgmental probing should be limited. The participants should be stimulated to report on their insight with relatively unstructured response and SR should be used as the focus of the research.

Concerning validity and reliability of SR, Gass & Mackey (2000) point out that the time delayed between the events and the recall should be minimized to increase the validity. Recall procedures should be designed to be strongly related to the focus of the study. Most importantly, it must be ensured that prompts or questions do not alter the cognitive process during the event which is considered to be a potential key to the validity issue of SR. The authors also provide several recall support mechanisms including relationship to specific action, temporal relationship, participants' training, structured-procedural, stimulus for recall and initiation of questions of recall interaction. It should be noted that SR is an indirect method of gathering cognitive process; hence, the finding should be cautiously interpreted. The point related to researchers' familiarity of the natural context should be considered, particularly on the researchers' bias in research design and data collection procedures (Lyle, 2003).

SR methodology is used to investigate cognitive factors in second language research (Gass & Mackey, 2007). However, there is a growing number of studies that employed SR to determine perception of feedback in second language learning

context (Gass & Mackey, 2007 and Polio, Gass & Chapin, 2006). This includes the use of SR to enhance students' reflection in technology-integrated classroom (Edwards-Leis, 2007).

Mackey, Gass & McDonough (2000, cited in Gass & Mackey, 2007) employed SR to investigate ten ESL learners' perceptions about the interactional feedback focusing on morphosyntax, phonology, lexi and semantics through picture-differences tasks. When the learners completed each task, they were shown with the recorded videotape during task performance and were asked to report on their thoughts during the interaction. The authors concluded that SR allowed the researchers to find learners' perceptions on the interactional feedback. Similarly, SR was significant that the data gained from this method can confirm the finding. SR was used to investigate the interactional patterns of non-native speakers interacting with second language learners outside classroom in which videotape stimuli were used to elicit the comments of the interaction.

However, very limited studies in language testing employed SR methodology to investigate strategies used while test takers were performing the test tasks, particularly in the speaking test. A study was conducted by Cohen & Olshtain (1993, cited in Cohen, 1998) on strategies in producing speech acts. The study investigated the ways that 15 non-native speakers assessed, planned and delivered speech acts. Six speech act situations which were two apologies, two complaints and two requests were given to the participants on written cards. They were read aloud to the participants twice at a time in English, and participants were called to participate in the role-play with appropriate responses to the situations. The interactions were videotaped and played back to them to stimulate the recall in their native language with both fixed and probing questions on the factors contributing to oral production in their responses. The retrospective verbal reports were analyzed on the processing strategies in their response formulation to the situations. The finding showed that general assessment of the utterance was employed half of the time without planning for specific grammar and vocabulary and often in two languages and sometimes in three languages with the utilization of different strategies to search for language forms. They did not use the strategies much for grammar and pronunciation. The

participants were characterized as metacognizers, avoiders and pragmatists from their speech production styles. Similar to the direct speaking test, sole and first study on a semi-direct internet-based speaking test by Swain et al. (2009) for the standardized TOEFL iBT was found to employ SR methodology in test taking strategy elicitation. Conforming to the recommendation from Gass & Mackey (2007), the participants in the study were trained to use SR. During the training session, they were encouraged to verbalize freely to the recall from the video stimulus. The instructions were translated and clarified into the participants' native language which was Chinese to ensure the thorough understanding. The exit interview was arranged later and the questions were in both English and Chinese. In the main study, 30 participants were encouraged to use both English and Chinese during the stimulated recall session when they completed each test task. They were asked to report about what they did before, during and after each speaking task. The researchers reminded the participants to report only on what they were thinking while they were performing the test task not on what they thought they should do or could have done. Two weeks later, the semi-structured exit interview was conducted to all participants to address the final thoughts. The authors claimed that SR methodology could raise test takers' awareness on strategies used in a speaking task and elicit the strategies they could add to their repertoire in order to choose the combination of effective strategies for given tasks. One limitation of SR methodology was the partial coverage of possible strategies that the test takers employed. Another point was on the pause between each task to elicit test takers' recalls that were likely to affect the reported strategies and test performance on the sub-sequent part. However, the authors argued that the latter limitation could be viewed as "the value of stimulated recall" (Swain et al., 2009:53). SR could help test takers in understanding the particular strategy that assisted them in the test taking context in specific task types and language skills.

## **2.8 Summary**

This chapter presents the review of related literature covering language for specific purposes (LSP) speaking ability's theory, studies on LSP language ability, LSP speaking tests, web-based language testing, advantages and limitations of web-based language testing, the effects of speaking test tasks on speaking performances,

test takers' attitudes towards the web-based and computer-based tests, test taking strategies and studies on the stimulated recall methodology in the test taking strategies. They are used in the research instruments development, data collection and analysis and discussions of the findings. The research methodology will be presented in the next chapter.

## **CHAPTER III**

### **RESEARCH METHODOLOGY**

This section presents the research methodology used in the study, covering research design and approach, population and sample, research instruments, data collection and data analysis.

#### **3.1 Research design and approach**

The following research design and approach were employed to answer the three research questions.

1. To what extent do the WBST-EFT task types and proficiency levels affect the speaking performances of high and low proficiency students in the English for Tourism course in terms of language and content knowledge, and the fluency of their speech performances?
2. What are the students' attitudes towards the WBST-EFT?
3. Are there any differences in the types and frequency of speaking test taking strategies used by high and low proficiency students in doing the WBST-EFT?

The research design in this study was 2x3 factorial design in which stratified sampling was used. Factorial design was selected due to the prominent characteristic that allowed the effects of multivariable to be investigated in one experimental study (Campbell & Stanley, 1966). In the present study, the effects of both the three task types and the two proficiency levels were investigated. The main and interaction effects of the two variables on the English for Tourism II speaking performances were examined.

#### **3.2 Population and sample**

The population was 230 third year students who took English for Tourism II course in the second semester of the academic year 2010 at Nakhon Ratchasima Rajabhat University (NRRU). Their majors were English and Tourism Industry. Their

faculties were Humanities and Social Sciences, Education and Management Science. They all took English for Tourism I as the pre-requisite course. For Tourism–major students, they can use the course grade as part of the application to get the bronze card tour guide license after they have completed all the requirements.

The sample group was 120 subjects who had English for Tourism I (EFT I) course grade of A, B+, B and D; and they were randomly selected to participate in this study. Then, they were classified into two groups based on their EFT I course grade. The mean score of the EFT I course grade was 67.50 and the standard deviations were 10.16. The Z value was calculated to categorize the subjects into the high and low proficiency groups. The 60 subjects who had the highest Z scores were assigned into the high proficiency group ( $Z=0.3$  to  $1.5$ ) and the 60 subjects with the lowest Z scores were categorized as the low proficiency group ( $Z=-0.5$  to  $-1.7$ ). Then, they were further divided into three sub-groups. Each sub-group consisted of 20 subjects and they were randomly assigned into three task types. Figure 3.1 illustrates the procedures of group assignment and the research design.

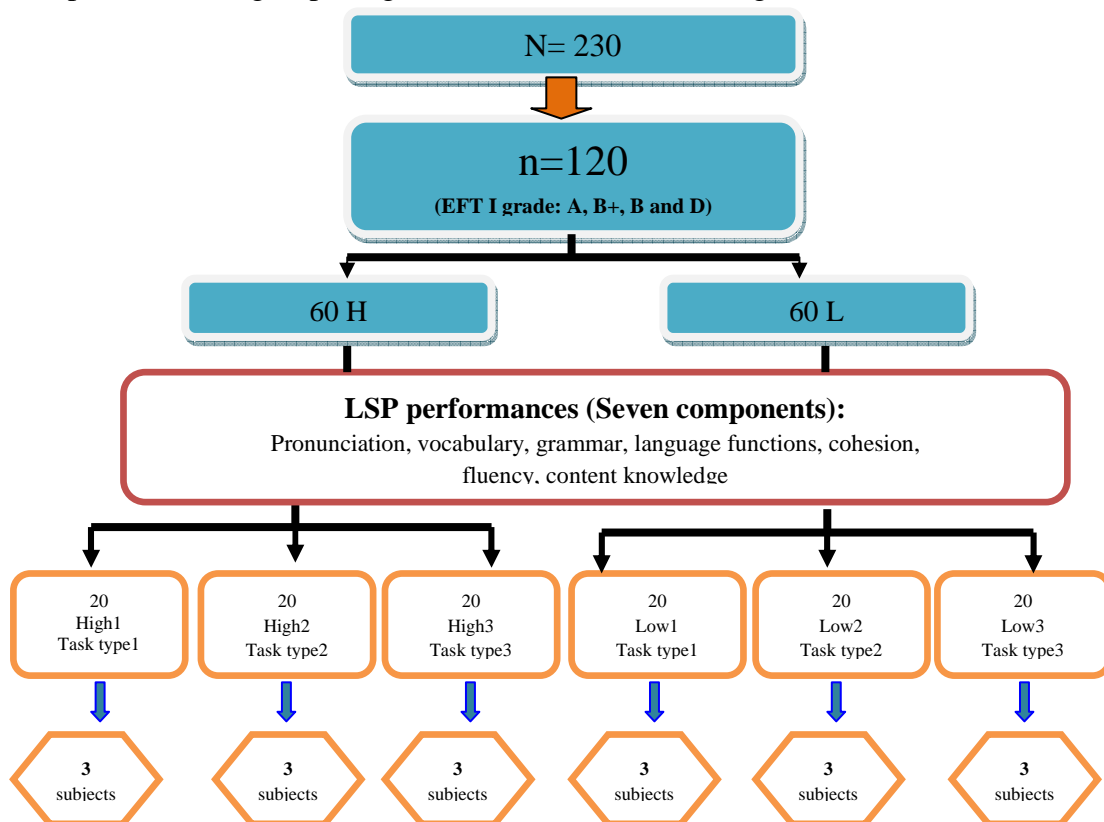


Figure 3.1: Group assignment with 2x3 factorial design using stratified sampling

From Figure 3.1, the six groups were required to take three different task types of the WBST-EFT. High1 and Low1 took Task type one, which focused on the presentation of cultural, historical and tourism information. High2 and Low2 took Task type two that was related to asking tourists with polite requests. High3 and Low3 took Task type three that dealt with enquiries and complaints. After that, the WBST-EFT online questionnaires were administered to all subjects to gain their perceptions towards this innovative test after the subjects had completed the test. At the end of each individual sub-task, three subjects from each group (18 in total) were randomly selected to participate in the stimulated recall interview session that aimed to elicit the types and frequency of strategies used in each task type.

### **3.3 Research stages**

The present study consisted of four stages. The first stage involved the analysis of both the specific purpose language use situations and the English for Tourism II course content. This stage included the selection of the language use tasks in tourism-related contexts. To obtain the language use tasks that closely corresponded to the real world tasks, the opinions from the subject specialists in the fields were gathered from the needs analysis questionnaire. The second stage was the development and validation of the research instruments which included the WBST-EFT with the rating scale, the attitudes towards the WBST-EFT online questionnaire, and the speaking test taking strategy interview. All of the instruments were validated by three experts in the field and revised before the next stage. A priori validity evidence on the content and construct of the WBST-EFT with the rating scale and the two instruments was obtained in this stage. The third stage was a pilot study of the instruments. The WBST-EFT was administered to 30 third year NRRU students in the first semester of the academic year 2010. These subjects completed EFT I as a pre-requisite course and they would take EFT II in the second semester which would be after the study. Therefore, the tutorial session for the EFT II was arranged for them due to the time constraint. After they had completed the WBST-EFT, the attitudes towards the WBST-EFT online questionnaires were administered to all subjects to obtain their reflections on this innovative speaking test. One subject from each group was required to participate in the stimulated recall interview session to elicit the



strategies used. The posteriori evidence on scoring validity, index of item facility (IF), item discrimination value (ID) and the inter-rater reliability value was established. All the instruments were revised before the main study. The revised version of the WBST-EFT was administered to 120 subjects following the similar steps as in the pilot study.

### **3.4 Research instruments**

#### **3.4.1 Web-based Speaking Test in English for Tourism (WBST-EFT)**

The WBST-EFT was developed under the theoretical framework of the LSP test development proposed by Douglas (2000) which has been modified from the framework of Bachman and Palmer (1996) in line with the interface design framework for technology-integrated test from Fulcher (2003b). These frameworks were integrated into the test development and the following figure illustrates the framework of the WBST-EFT. This figure was modified from the test development of Bachman and Palmer (1996:87).

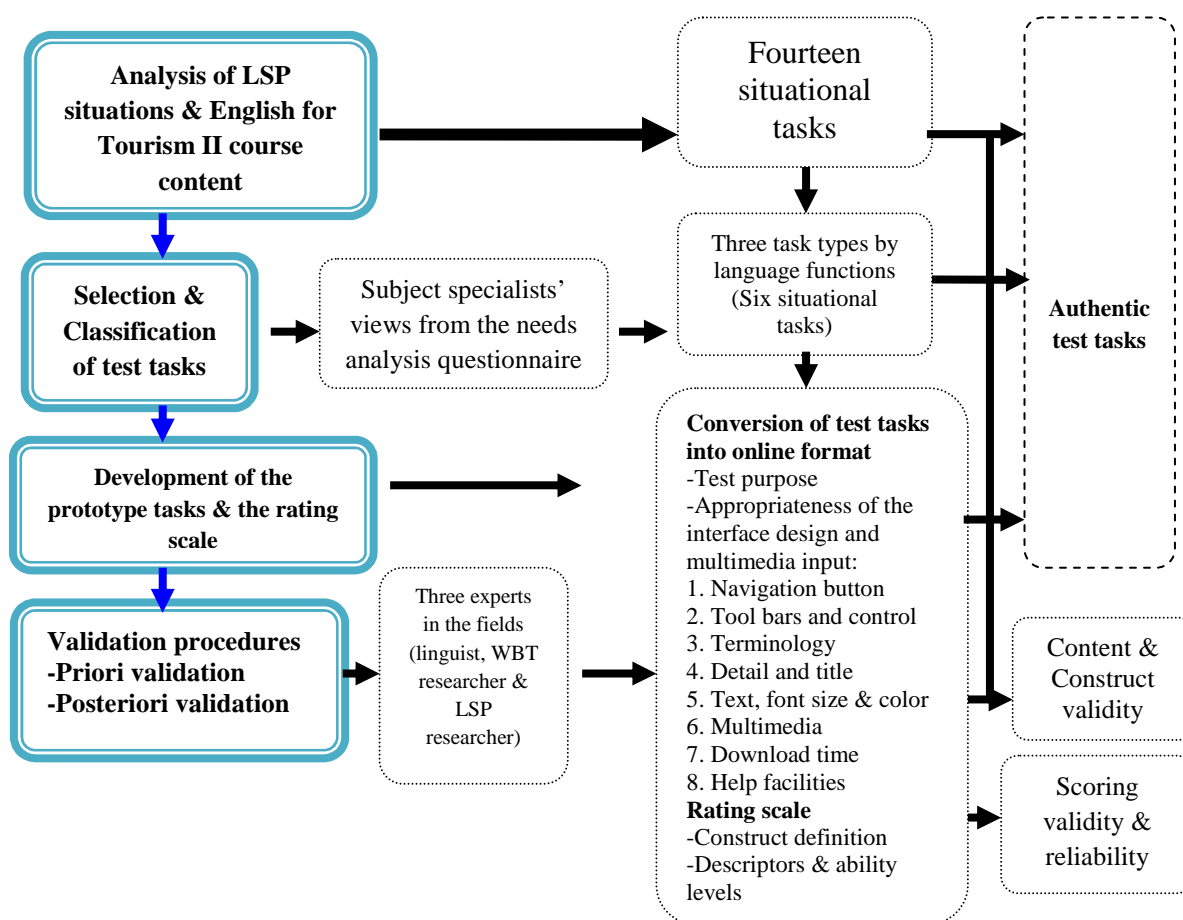


Figure 3.2: The WBST-EFT framework

### 3.4.1.1 Analysis of specific purpose language use situations and English for Tourism II course content

The initial step in the WBST-EFT development was the analysis of specific purpose language use situations and the selection of the language use tasks in tourism-related contexts (Douglas, 2000). This step involved conducting the needs analysis with the experts in the field with a needs analysis questionnaire (See Appendix A). Additionally, the WBST-EFT was the final achievement test in English for Tourism II (EFT II) course; hence, the information on the needs analysis questionnaire was gathered from this LSP course content analysis and the review of the related literature.

The analysis from the EFT II course content and related literature showed that there were 14 situational tasks that were taught in this course and frequently performed by the tour guides. The analysis also showed the course grading criteria

which were on accuracy, range, fluency and appropriateness of the language production. Due to the objective of the WBST-EFT as a final achievement test in EFT II course, all of this information was incorporated in the needs analysis questionnaire to obtain the test tasks that could both represent the salient characteristics of an LSP test from the experts in the field, and also meet the course requirement.

Concerning the needs analysis questionnaire, it was developed by the researcher. The objective of this instrument was to investigate the target language use tasks and situations, language knowledge required in professional tour guides, and criteria for assessing the language knowledge from the subject specialists in the fields. The information obtained from this instrument was used in the development of the WBST-EFT consisting of the test tasks that closely corresponded to the real world tasks and the rating scale to evaluate actual language knowledge used by professional tour guides.

There were four parts in the questionnaire. Part One was related to the demographic information of the content area specialists. Part Two and Part Three provided a 4-point Likert scale asking the degree of importance about the tasks, situations and language knowledge which were most likely to be used by professional tour guides. In addition, Part Three touched on additional language knowledge and testing. Part Four was also a 4-point Likert scale with open-ended questions that asked about the appropriateness of the criteria for assessing the language knowledge of the tour guides.

The questionnaire was validated by three experts in the field with the index of item-objective congruence (IOC) value at 1.00 for each item. Some parts of the questionnaire were revised according to the experts' suggestions and tried out with three subject specialists. Then, the instrument was revised again, particularly on the clarity of language in some parts. The final questionnaire was administered to 15 subject specialists who each had a minimum of 7 years of experience in tourism and English for tourism instruction. This group of people included five travel agency owners who used tour guides, five English-speaking professional tour guides who held bronze type tour guide license, and five English for Tourism II course lecturers.

To ensure the validity of the information, they were required to provide demographic information which was directly related to their expertise in Tourism or in English. Almost all of them had master's degrees in tourism or in English and one had a bachelor's degree in tourism. Their experience in their profession ranged from 7 to 9 years, except for two of the specialists who had 35 to 37 years in the field.

Results from a needs analysis questionnaire were based on the degree of importance from very important (4) to not important (1), and are presented in the following section.

#### **3.4.1.2 Selection and classification of tasks and situations**

The next step was the selection and classification of tasks and situations. The tasks that were rated by most of the specialists only in 'very important' and 'important' category were included in the WBST-EFT. From the 14 situational tasks, they were grouped into six situational tasks. These tasks were further classified into three task types as follows.

Task type one: Presenting tourism-related information

Task type two: Giving polite suggestion to the tourists about what they should do and should not do in Thailand

Task type three: Dealing with tourists' enquiries and complaints

Each task type was made up of two sub-tasks and there were six sub-tasks in total. In addition, the subject specialists were asked to specify the degree of importance for the components of language knowledge used by the tour guides when organizing trips. These components were also proposed in Douglas's (2000) LSP ability framework in line with Fulcher's (2003a) speaking ability model. The components that were pertinent to English for Tourism II course syllabus were included. The information obtained in this part was used in the rating scale construction.

The results of the questionnaire revealed that the knowledge of vocabulary was rated as a very important feature by almost all of the subject specialists, and

fluency and content knowledge also posed the same degree of importance. There was consensus by the specialists that knowledge of grammar, pronunciation, language functions and cohesion were also considered as important features in the tourism domain.

With regards to the appropriateness of the criteria for assessing language knowledge, all of the specialists agreed that the range of speech was the most appropriate criterion. Appropriate use of grammatical structures and language functions with the consideration of sociolinguistic domains was also rated as an important criterion, whereas both accuracy and fluency were rated as very appropriate. Therefore, all these criteria were included in the rating scale.

It is clear at this stage that the analysis of the specific characteristics of the target language used in context and tasks is a vital procedure in the WBST-EFT development. Without this procedure, the WBST-EFT will not cover the important elements of the actual tasks in tourism context and it will directly affect the authenticity of the test tasks.

Regarding the specificity of LSP tests, Douglas (2001) mentions that all tests are developed for some purposes and they will fit in the particular point of the continuum of specificity. LSP tests therefore must include certain and precise characteristics used by people in the profession such as specific pronunciation, vocabulary, word meaning, and sentence structures. People who are not in the fields will not have a thorough understanding of these characteristics. Thus, the subject specialists' view was purposively obtained in this study in order to provide the specific features that must be included in the WBST-EFT tasks and the rating scale.

#### **3.4.1.3 Development of the WBST-EFT prototype tasks and the rating scale**

Another step dealt with the test specification and the actual test tasks development. The test specification or test blueprint was written to be used as a plan for the WBST-EFT construction. This is an important step that cannot be excluded from any test developments. This planning guided the WBST-EFT test developer

about the test purposes and language ability to be measured. The blueprint was also used as the guideline for the WBST-EFT writing and task construction item. In addition, it also gave details on the scoring criteria, procedures and interpretation for the raters. It provided information on test objectives, test construct and interpretation of test performances for the test users. Finally, the details of the test specification can be used as part of the validation procedure to provide empirical evidence on test validity.

Concerning the test construct, Douglas's (2000) LSP ability framework and Fulcher's (2003a) speaking ability were incorporated in the WBST-EFT. Some components of LSP language ability were selected on their relatedness to English for Tourism II course. As a part of LSP construct, background knowledge was investigated by some studies with varied results (Clapham, 1996 cited in Douglas, 2000 and Krekeler, 2006). However, some studies revealed the supportive effect of this knowledge on test performance (Clapham, 1996 cited in Douglas, 2000); therefore, it was included in the WBST-EFT. Fluency of speaking ability was used as a criterion in several LSP speaking tests (Brown, 1995, ILEC Handbook, 2008, BULATS Handbook, 2009); thus, it was also included in the WBST-EFT construct.

After the specification had been drafted, the actual test tasks were constructed and converted into the online format. In the WBST-EFT, the students would act as the tour guides organizing the trip in the central part of Thailand. Drawing on previous studies, English for Tourism II course syllabus analysis, and data derived from the needs analysis questionnaire, the researcher created three target language use task types and six situational tasks. These types of language and situations were most likely to be used by professional tour guides. There were three sections in the WBST-EFT, which were categorized by task types, and each task type had two sub-tasks. Each test task purposively incorporated multimedia in order to simulate a real world task and make it live. The students had preparation time in Sections One and Two. The information about the preparation and response time and marking criteria was available in the instruction part of the test. The whole test lasted approximately 24 minutes. The students could take the sample test so as to be familiar with the test. They were required to respond to each test task by clicking on the record button and

start speaking through the microphone when they heard the sound ‘beep’ and saw the ‘start speaking’ prompt on the screen. After they finished speaking, they had to click on the same button again to stop. They were allowed to record their responses only once. Then, they clicked the next button to move onto the next task.

In Sections One and Two, pictures with audio input were presented to elicit the ideational and manipulative functions of the students. In Section Three, short video clips which simulated real world scenarios were presented. The last part was reciprocal in nature requiring the students to interact with the scenarios by using the heuristic language function.

In Section Three, after the scenarios that were likely to happen in the organized tour were selected, short video clips were created. This section aimed to elicit students’ heuristic function. The TLU characteristics were used in creating the dialogues between the tour guide and the tourists. There were six scenarios in this section and three of them were classified into responding to tourists’ enquiries, and the other three dealt with complaints. All of the topics and situations were selected from previous studies and from the suggestions of the subject specialists. The dialogues were checked by a linguist and a professional tour guide. Some parts were revised to improve the clarity and accuracy of language, particularly at the discourse level.

The test tasks were then converted into short films. Eight young ambassadors at Nakhon Ratchasima Rajabhat University (NRRU) voluntarily participated in this research project. This group of students was trained to be professional tour guides. There was a rehearsal of all six scenarios before the actual filming took place. To increase the authenticity of test tasks, all of the scenes were recorded at actual sites such as at the hotel, on the bus, and at the tourist attractions. The films were edited by the researcher using free downloadable software programs: Window Movie Maker version 2.6 and Sound Forge Trial version 9.0.

All of the six prototype tasks were posted on Moodle version 1.9.5, a free online platform that is currently being used at NRRU. The students’ speech productions were recorded with Sound Forge software program version 9. Their

responses were stored in this platform and can be retrieved online by raters. The details and objectives of the test tasks are presented below.

#### Section One (Task type one): Presenting tourism-related information

This section aimed to elicit the students' ability in presenting national tourist attractions and explaining the tour program.

##### Task One: Presenting tourist attractions

The first task aimed at eliciting the students' ability to present two of the most famous national attractions in Thailand: the Emerald Buddha Temple and the Grand Palace. The students were provided with seven pictures about the two sites (four about the Emerald Buddha Temple and three about the Grand Palace) and they were asked to explain these pictures in details. They had 7 minutes to work on this task. For each picture, they had 20 seconds to prepare their responses and the remaining 40 seconds were for their speech production.

##### Task Two: Describing one-day tour program in the central region of Thailand

In Task Two, the students first read the one-day tour itinerary, and then they were required to present the information to tourists. They were asked to provide additional details of the underlined attractions.

#### Section Two (Task type two): Giving polite suggestions to tourists

The objective of Task type two was to assess the students' ability in giving polite suggestions to tourists in two different situations.

##### Task Three: At the Summer Palace

The students first watched the video clip containing a monologue of the tour guide at the Summer Palace. Then, there were six pictures on the clip which required the students to give polite suggestions on what the tourists should do and should not do in each situation based on Thai cultural and religious beliefs.



#### Task Four: At Jatujak Market

The students attempted this task in a similar function to Task Three. They first watched a video clip and were asked to respond to the six pictures containing different scenes by giving polite suggestions regarding what the tourists should do at the crowded shopping center.

#### Section Three (Task type three): Dealing with enquiries and complaints

Task type three emphasized the students' ability to deal with tourists' enquiries and complaints on a variety of topics.

#### Task Five: Dealing with enquiries

The students first watched the video clip containing the dialogue of three different enquiries: asking for help in recovering a stolen wallet, requesting a guide to explore the night life, and requesting medical assistance. At the end of each dialogue, the students were asked to respond to the enquiry politely and appropriately.

#### Task Six: Dealing with complaints

Task Six incorporated three complaints: an incomplete tour program, an unrequested hotel room, and prolonged wait for a bus. The students first watched video clips containing different complaints, and they were required to respond to each complaint politely and appropriately (See Appendix B for the WBST-EFT).

- **The rating scale construction**

The WBST-EFT rating scale was an essential instrument in scoring the speaking performances. It was specifically used with the target population and test purpose. The WBST-EFT rating scale provided operational definitions of LSP construct in tourism and levels of mastery of these features in completing the test tasks. The description in the scale was explicit, precise and able to differentiate the students' levels of mastery of the construct. Raters needed to be trained to use the scoring scale in order to obtain the reliability of their rating. As with the test, the rating scale was designed under the theoretical frameworks and underwent the

validation process. The WBST-EFT rating scale was developed under the rating scale development proposed by Fulcher (2003a). The following details of the development procedure are presented to justify the use of this instrument in scoring procedure and interpretation.

First, the language ability in the rating scale was defined. As part of the WBST-EFT, the rating scale employed similar construct of the LSP ability proposed by Douglas (2000) and speaking ability framework proposed by Fulcher (2003a). Then, the purpose and type of rating scale were decided. It was assumed that both the purpose and type of rating scale could guide the rating procedures and scoring interpretation. Thus, they were purposively decided with the consideration of their usefulness and suitability with the WBST-EFT. The purpose of the WBST-EFT rating scale was to guide the rating process emphasizing the quality of the performance. Therefore, it was considered as the “Assessor-oriented scale” (Fulcher, 2003a). The scale contained operational construct definitions that were easy to comprehend within a short time.

Regarding the type of scale, the analytic rating scale was used in this study due to its appropriateness with the test purpose as the classroom final achievement test and for the diagnostic purpose. The analytical rating scale allowed for assessing specific components of language ability defined by the construct definitions. Additionally, each of the scale descriptors contained a specific level of mastery of language ability. Therefore, either the mastery or failure of the specific language components could be indicated. This analytic scale could provide information related to the strengths and weaknesses of the students, and the information could be used for remedial courses and instructional approach designs.

In terms of criteria for correctness, accuracy was used in rating the speaking responses. The notions of accuracy in linguistic elements, range, complexity and appropriateness of speech production were used as the criteria by a number of LSP speaking tests (Brown, 1995, ILEC Handbook, 2008, BULATS Handbook, 2009). Therefore, these elements were included in the WBST-EFT rating scale.

After that, the number of levels of ability on the scale was decided and the band descriptors were written. Like many tests, the rating scale required an appropriate design to derive the number of levels of ability and descriptors that are clear and precise in differentiating the mastery of the students' language ability in completing the test tasks. Approaches to a rating scale design proposed by Fulcher (2003a) were used to develop the descriptors of the WBST-EFT rating scale. The design of the scale was based on "the expert-judgement method". The researcher who taught the English for Tourism II course for 4 years and worked as a professional tour guide for 7 years wrote the band descriptors. The expert-judgement method in rating scale design requires a number of years in field experience (Fulcher, 2003a). Prior to band descriptors development, the number of levels of ability was decided. The WBST-EFT consisted of five language abilities starting from level 0 (a very poor user) to 4 (a very good user). The number of ability levels was pertinent to the course grades which started from F and went up to A. In this way, the band levels were conveniently converted into course grades. The sequence of band descriptors meaningfully and clearly reflected a progression in LSP language ability. The band descriptors were relevant to the language requirements stated in the course syllabus, and were based on the experts' recommendations from the needs analysis questionnaire. The LSP ability and speaking theoretical framework were included in the rating scale. The criteria for correctness were modified from Fulcher's (2003a) speaking ability (See Appendix C for the rating scale).

Before rater training, rater selection was conducted to ensure that all raters were qualified. The final step was rater training to obtain inter-rater consistency and to assure the reliability of the test scores. All the raters were provided with the rating form, descriptors of the criteria and descriptions of rating procedures. They were trained to understand the descriptors and criteria, to follow the rating procedures and to appropriately apply the rating scale. This session was carefully arranged because it could affect the scoring validity and reliability of the test.

#### **3.4.1.4 Validation procedures**

In validating the WBST-EFT, both priori and posteriori validation procedures were adopted. To establish the content and construct validity evidence, the two instruments were validated by three experts in the field by using the index of item-objective congruence ( $IOC > .75$ ). The description of the test specification was used to establish its validity evidence for both content and construct (Weir, 2005 & Bachman & Palmer, 1996). The priori validity evidence on both the content and construct revealed the IOC value of 1.00 for each test task and descriptor in the rating scale. This reflected a high validity of the content and construct of the instruments. As for clarity of the language, the test and rating scale were revised according to the experts' suggestions.

After the two instruments were validated, they were piloted with 30 subjects in the first semester in the academic year 2010 and these subjects were excluded from the main study. They were classified into two groups based on the Z scores and they were further divided into three sub-groups. Each sub-group consisted of five subjects. In each proficiency level, the subjects were randomly assigned to three task type groups. As for the posteriori validity evidence regarding scoring validity and reliability, item analysis of the WBST-EFT was carried out. The result yielded high values of item discrimination index which ranged from .58 to .63 for the six tasks. This means all the test tasks could effectively differentiate the mastery levels of all the subjects. For the difficulty level of the test tasks, the values ranged from .32 to .35 which can be interpreted that the test was quite difficult. Pearson Correlation was applied to assess the inter-rater reliability of the rating scale. The correlation coefficient was .70 and Cronbach's alpha value was .98 which reflected a high reliability of rating and could be claimed that raters were highly consistent in their rating.

#### **3.4.2 The Attitudes towards the WBST-EFT Online Questionnaire**

The instrument was developed to investigate the students' perceptions towards the test on the overall usefulness of the test, appropriateness of time for preparation and response formulation, task difficulty and the interface design. This instrument

was developed through stages similar to those for the needs analysis questionnaire. It was validated by three experts in the field using the index of item-objective congruence (IOC) and all of the items posed the IOC value of 1.00 reflecting high content and construct validity. Some parts of the questionnaire were excluded based on the experts' suggestions, particularly on the unrelated information including students' age and years of exposure to English. After that the questionnaire was converted into the online version using the freeware program Google document and posted on the WBST-EFT. Then, it was tried out in the pilot study with 30 subjects in the first semester of the academic year 2010 and revised before the main study. The data obtained from this questionnaire in the pilot study was also included in the test revision. It took approximately 10 to 15 minutes to complete the questionnaire. The questionnaire consisted of three main parts: demographic information, opinions towards the WBST-EFT and open-ended questions. The first part asked for demographic information to collect the general information concerning the subjects' faculty, English for Tourism I course grade and the experience in any computer-based or web-based test taking. The second part was the students' opinions towards the WBST-EFT. The 4-point Likert scale was used to gain the subjects' reflections on the test. This part incorporated four main topics. The first topic was the overall usefulness of the test, asking the subjects to demonstrate both their strengths and weaknesses in speaking ability, accuracy and fairness in rating procedures and tasks, clarity of instructions and tasks, appropriateness of the test tasks and test taking procedures and test taking anxiety. Some of the questions on the attitudes on computer-based test taking were adapted from the studies of Kenyon and Malabonga (2001) and Norris (2001). The second topic was the appropriateness of time for preparation and response formulation, and the third topic was the task difficulty. The last topic was the appropriateness of the interface design and the questions were from Fulcher's (2003b) framework. The final part contained two open-ended questions on the strengths and weaknesses of the WBST-EFT and what the subjects liked and did not like most about the test. The last part aimed at gaining more in-depth views from the subjects (See Appendix D for the attitudes towards the WBST-EFT online questionnaire).

### **3.4.3 The Speaking Test Taking Strategies Interview**

A semi-structured interview on speaking test taking strategies was constructed to elicit the strategies used in taking the WBST-EFT (See Appendix E for the interview questions). Before writing the interview questions, the list of strategies used were developed to provide the description and classification of strategies. The list was used later in the coding scheme to exemplify the actual strategies used during the data analysis procedure. The list was derived from the compilation of the related literature on speaking strategies, L2 use and learning, test taking and communication strategies and internet-based speaking test taking strategies (Fulcher, 2003a, Bachman & Palmer, 1996, Cohen, 1998 and Swain et al., 2009). It consisted of three main types of strategies: Communication strategy (achievement and avoidance), Cognitive strategy (selecting, comprehending, storing memory and retrieval), and Metacognitive strategy (goal setting, organizing, planning and evaluating). After the list had been developed, the interview questions were drafted. These interview questions were employed to investigate the types and frequency of test taking strategies used in completing the WBST-EFT. The interview included the scripts and questions used during the stimulated recall session to elicit the test taking strategies. Similar to other instruments, the list of strategies used and the interview questions were validated by three experts in the field using the IOC table. All of the items on the list, scripts and interview questions had the IOC value of 1.00. Similar to the previous instruments, the instructions on the interview questions and scripts were revised for the clarity of language based on the experts' suggestions before the pilot and main study were conducted. The reported strategies were transcribed and categorized by two coders to avoid inconsistency and bias on the data. One of the coders was the researcher and the other one was an English lecturer who had M.A.in Applied Linguistics with 35 years of experience in the profession.

### **3.5 Data collection**

After the pilot study, the main study was conducted and replicated the stages of the pilot study. The revised version of the WBST-EFT and the rating scale was administered to 120 subjects who were the third year students from NRRU in the second semester of the academic year 2010. They were classified into six groups as explained in the group assignment procedures illustrated in Figure 3.1 (page 67) and each group of subjects was assigned to do different task types. Their performances were audio recorded and stored in the database to be rated later by two raters.

After the subjects had completed the WBST-EFT, they were required to respond to the WBST-EFT attitudes online questionnaire. Eighteen subjects who were randomly selected from the six groups participated in the stimulated recall interview session on strategies used report before they completed the online questionnaire.

In the stimulated recall interview session, eighteen subjects were asked to watch the playback video clips of their recorded performances when they attempted each test task. They were asked to report on their thought before, during, and after the task performance. They could pause the clips any time they wanted to while giving their report on their strategies used. Their verbal report responses were transcribed and coded by two coders. Similar to the rating procedure, the coders were trained to understand the strategy taxonomy and coding procedure before the pilot study.

### **3.6 Rating for test scores**

The rating for test scores was conducted by two raters. The criteria for rater selection were from the years of experience in their profession and their language proficiency scores. One of the two raters was the researcher and the other was an English for Tourism II course lecturer who had 37 years of teaching experience and was also a trainer of the TAT tour guide training course. After the pilot study, two raters tried out the rating scale with ten sample speeches. When there was any discrepancy of the band score between two raters, the discussion was arranged in order to come up with the mutual agreement based on the scale. Some parts of the

descriptors were revised before being used in the main study. The sample speeches with the band scores are presented in Appendix F.

### **3.7 Data analysis**

The data analysis in the present study was conducted following the quantitative and qualitative approaches.

1. To answer the first research question “To what extent do the WBST-EFT task types and proficiency levels affect the speaking performances of high and low proficiency students in the English for Tourism course in terms of language and content knowledge, and the fluency of their speech performances?”, descriptive statistics and the two-way ANOVA were carried out to check for the significant differences among the mean scores from the two proficiency groups on the three task types. If the value was significant, this indicated the existence of at least one significant difference between group means. In addition, a post-hoc Scheffé test was performed to indicate the significance of the particular contrast. Content analysis on speaking performances was conducted to further investigate the extent that each element of language and content knowledge, and fluency were affected by the proficiency levels and three task types.

2. To answer the second research question “What are the students’ attitudes towards the WBST-EFT?”, independent samples t-test was calculated on the students’ perceptions towards the test from the scores of the online questionnaire. Content analysis was conducted with the open-ended questions to categorize the responses.

3. To answer the third research question “Are there any differences in types and frequency of speaking test taking strategies used by high and low proficiency students in doing the WBST-EFT?”, a qualitative content analysis was employed to transcribe and code the data to investigate whether there were any differences and similarities in types of strategies used by the two proficiency groups for each task type. Frequency of responses on the types of reported strategies was also calculated. Percentage concerning the total number of strategies reported for the whole test and in relation to each strategy category was calculated.



### **3.8 Summary**

In summary, this chapter provides the research methodology, research design and approach, population and sample, research instruments, data collection and data analysis. The results of the main study are presented in the next chapter.

## CHAPTER IV

### RESULTS

This chapter presents the results of the main study in relation to the three research questions.

**4.1 Research question 1:** To what extent do the WBST-EFT task types and proficiency levels affect the speaking performances of high and low proficiency students in the English for Tourism course in terms of language and content knowledge, and the fluency of their speech performances?

Research hypotheses:

- 4.1.1. There is no significant difference between the mean score of the high proficiency students and that of the low proficiency students at the .05 level.
- 4.1.2. There are no significant differences at the .05 level in the three task types performed by the two proficiency groups.
- 4.1.3. There are no significant interaction effects at the .05 level between the two proficiency groups and the three task types.

For this research question, the two-way ANOVA was conducted to investigate both the main and interaction effects between the proficiency levels and the task types on students' speaking performances. The Statistical Package for Social Sciences (SPSS) version 13.5 was employed to analyze the data. The two-way ANOVA results of the effects from the two variables on the total speaking scores are displayed in Table 4.1. The mean scores and standard deviations of the two proficiency groups in attempting the three task types are also presented.

#### 4.1.1 The effects of proficiency levels and task types on the total LSP speaking scores

**Table 4.1: The effects of proficiency levels and task types on LSP speaking scores using 2x3 ANOVA**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	14.13(a)	5	2.83	12.06	.00	
Intercept	635.50	1	635.50	2712.65	.00	
Proficiency levels	12.90	1	12.90	55.02	.00*	0.32
Task types	1.18	2	.59	2.53	.08	
Proficiency levels * Task types	.056	2	.03	.12	.89	
Error	26.71	114	.23			
<b>Total</b>	<b>676.33</b>	<b>120</b>				
<b>Corrected Total</b>	<b>40.84</b>	<b>119</b>				

\* $p \leq .05$

In Table 4.1, the results from the two-way ANOVA show that the only statistical significance at .05 level occurs in the proficiency levels of the two proficiency groups' total scores,  $F(1, 114) = 55.02$ ,  $p < .05$  with only a small effect size ( $\eta^2 = .32$ ). This indicates that only proficiency levels had a main effect on the LSP performances. This means that the two proficiency groups' performances were significantly different. Hence, the first hypothesis is rejected. However, the statistical result does not show any statistical significance among the total scores of the three task types,  $F(2, 114) = 2.53$ ,  $p > .05$ . Similar to the task types, there is no statistical significance between the two proficiency groups and the three task types on the total scores,  $F(2, 114) = .12$ ,  $p > .05$ , showing that there was neither main effect from the task types nor interaction effect between the proficiency levels and the task types on the LSP performances. In other words, the two proficiency groups' performances

were not significantly different across the three task types, so the second and third hypotheses are accepted. The following table shows the means and standard deviations of the scores of the six groups.

**Table 4.2: Descriptive statistics of the speaking scores of the six groups**

Proficiency levels	Task types	Mean	SD
<b>High</b>	1.00	2.78	.31
	2.00	2.48	.39
	3.00	2.62	.69
	<b>Total</b>	<b>2.63</b>	<b>.50</b>
<b>Low</b>	1.00	2.06	.28
	2.00	1.87	.56
	3.00	1.98	.52
	<b>Total</b>	<b>1.97</b>	<b>.47</b>
<b>Total</b>	1.00	2.42	.46
	2.00	2.18	.57
	3.00	2.30	.68
	<b>Total</b>	<b>2.30</b>	<b>.58</b>

Table 4.2 shows that the total mean score of the high proficiency group is more than that of the low proficiency group ( $\bar{x}_H=2.63$ ,  $\bar{x}_L=1.97$ ). The variations among scores in the two proficiency groups are not much and range from .28 to .69. In other words, the high proficiency group performed better than the low proficiency group. Focusing on the total scores of the three Task types, Task type one has the highest mean whereas the lowest mean is in Task type two ( $\bar{x}_{T1}=2.42$ ,  $\bar{x}_{T2}=2.18$ ). Task type three poses the largest spreading of scores whereas the smallest one is in Task type one (Task3 SD=.68, Task 1 SD= .46). This means that the two proficiency groups had the best performances in Task type one while the worst performances were in Task type two. The following part displays the effects of proficiency levels and task types on the LSP individual component.

#### **4.1.2 The effects of proficiency levels and task types on the LSP individual component scores**

Although statistical significance was only found in the proficiency levels from the total scores, a two-way ANOVA test of the individual component of the speaking

performance was conducted to determine whether there were any significant main and interaction effects of the two variables. The speaking performance components consisted of pronunciation, vocabulary, grammar, language functions, cohesion, fluency and content knowledge. The following table presents the effects of proficiency levels and task types on pronunciation mean scores.

**Table 4.3: The effects of proficiency levels and task types on pronunciation scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	9.94(a)	5	1.98	6.68	.00	0.21
Intercept	632.50	1	632.50	2125.03	.00	
Proficiency levels	9.35	1	9.35	31.42	.00*	
Task types	.57	2	.28	.97	.38	
Proficiency levels * Task types	.01	2	.00	.02	.97	
Error	33.93	114	.29			
<b>Total</b>	<b>676.37</b>	<b>120</b>				
<b>Corrected Total</b>	<b>43.87</b>	<b>119</b>				

\* $p \leq .05$

Table 4.3 shows that only proficiency levels have a statistical significance at .05 level on the pronunciation scores of the two proficiency groups,  $F(1, 114) = 31.42$ ,  $p < .05$  with the effect size of .21. This is the only main effect on this component. This means that the pronunciation performances of the two proficiency groups were significantly different. The means and standard deviations of the pronunciation scores are provided in the following table.

**Table 4.4: Descriptive statistics of pronunciation scores**

<b>Proficiency levels</b>	<b>Task types</b>	<b>Mean</b>	<b>SD.</b>
<b>High</b>	1.00	2.51	.40
	2.00	2.52	.48
	3.00	2.68	.63
	<b>Total</b>	<b>2.57</b>	<b>.51</b>
<b>Low</b>	1.00	1.97	.49
	2.00	1.97	.57
	3.00	2.10	.64
	<b>Total</b>	<b>2.01</b>	<b>.56</b>
<b>Total 3 tasks</b>		<b>2.29</b>	<b>.60</b>

In Table 4.4, the high proficiency group has higher total means of pronunciation than the low proficiency group ( $\bar{x}_H=2.57$ ,  $\bar{x}_L=2.01$ ), showing that the high proficiency group performed better in pronunciation than the low proficiency group. The standard deviations show homogeneous variances among the scores. The comparison of the standard deviations shows that the low proficiency group poses similar spreading of scores (SD=.56) to that of the high proficiency group (SD=.51). This means that the variations of the scores from the two proficiency groups were not much different. The following table presents the effects of proficiency levels and task types on vocabulary mean scores.

**Table 4.5: The effects of proficiency levels and task types on vocabulary scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	18.04(a)	5	3.61	11.46	0	
Intercept	644.03	1	644.03	2044.05	0	
Proficiency levels	14.7	1	14.7	46.66	.00*	0.27
Task types	3.00	2	1.50	4.76	.01*	0.05
Proficiency levels * Task types	0.34	2	0.17	0.55	0.57	
Error	35.91	114	0.31			
<b>Total</b>	<b>698</b>	<b>120</b>				
<b>Corrected Total</b>	<b>53.96</b>	<b>119</b>				

\* $p \leq .05$ 

The results from Table 4.5 show a statistical significance at .05 level of both the proficiency levels,  $F(1, 114) = 46.66, p < .05$  and the task types,  $F(2, 114) = 4.76, p < .05$  on the vocabulary total scores, indicating that both variables had major effects on vocabulary performances. The effect size of the proficiency levels is .27 whereas that of the task types is only .05. This means that the two proficiency groups' vocabulary means were significantly different among the three task types. A Scheffé post-hoc test was conducted to compare the mean difference.

**Table 4.6: Scheffé post-hoc test on vocabulary scores**

(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.38 (*)	.12	.01	.06	.69
	3.00	.13	.12	.58	-.18	.44
2.00	1.00	-.38 (*)	.12	.01	-.69	-.06
	3.00	-.25	.12	.14	-.561	.06
3.00	1.00	-.13	.12	.58	-.442	.18
	2.00	.25	.12	.14	-.06	.56

\* $p \leq .05$ 

In Table 4.6, the mean difference between Task type one and Task type two is .38 ( $\bar{x}_{T1}=2.48$ ,  $\bar{x}_{T2}=2.10$ ). It is the only significant difference at .05 level of the vocabulary mean scores. This indicates that the vocabulary scores only differ between Task type one and two. The means and standard deviations of the vocabulary scores are given below.

**Table 4.7: Descriptive statistics of vocabulary scores**

Proficiency levels	Task types	Mean	SD
<b>High</b>	1.00	2.90	.31
	2.00	2.46	.45
	3.00	2.63	.75
	<b>Total</b>	<b>2.66</b>	<b>.56</b>
<b>Low</b>	1.00	2.07	.28
	2.00	1.75	.71
	3.00	2.07	.65
	<b>Total</b>	<b>1.96</b>	<b>.59</b>
<b>Total</b>	1.00	2.48	.51
	2.00	2.10	.69
	3.00	2.35	.75
<b>Total 3 tasks</b>		<b>2.31</b>	<b>.67</b>

From Table 4.7, the high proficiency group poses higher total of vocabulary mean scores than that of the low proficiency group ( $\bar{x}_H=2.66$ ,  $\bar{x}_L=1.96$ ), showing that



the high proficiency group had better vocabulary performances than the low proficiency group. It can also be seen that the most difference in the vocabulary mean scores between the high and low proficiency groups is in Task type one ( $\bar{x}_H=2.90$ ,  $\bar{x}_L=2.07$ ) whereas the least difference is in Task type three ( $\bar{x}_H=2.63$ ,  $\bar{x}_L=2.07$ ). This means that they had the most different vocabulary performance in Task type one, while the least different performance was in Task type three. Both proficiency groups have similar standard deviations (High SD=.56, Low SD=.59). The most difference in the spreading of the scores between the two proficiency groups is in Task type two (High SD=.45, Low SD=.71) while the least difference is in Task type one (High SD=.31, Low SD=.28). The following table illustrates the effects of proficiency levels and task types on grammar mean scores.

**Table 4.8: The effects of proficiency levels and task types on grammar scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	16.23(a)	5	3.24	10.56	0	0.28
Intercept	646.35	1	646.35	2103.00	0	
Proficiency levels	14.35	1	14.35	46.70	.00*	
Task types	1.26	2	0.63	2.06	0.13	
Proficiency levels * Task types	0.61	2	0.30	1.00	0.37	
Error	35.03	114	0.30			
<b>Total</b>	<b>697.62</b>	<b>120</b>				
<b>Corrected Total</b>	<b>51.27</b>	<b>119</b>				

\* $p \leq .05$

Table 4.8 illustrates that the only significant main effect at .05 level on grammar scores occurs in proficiency levels,  $F(1, 114) = 46.70$ ,  $p < .05$  with the effect size at .28. This means that the grammar performances of the two proficiency groups were significantly different. The descriptive statistics of the grammar scores are given below.

**Table 4.9: Descriptive statistics of grammar scores**

<b>Proficiency levels</b>	<b>Task types</b>	<b>Mean</b>	<b>SD</b>
<b>High</b>	1.00	2.87	.38
	2.00	2.45	.51
	3.00	2.67	.72
	<b>Total</b>	<b>2.66</b>	<b>.57</b>
<b>Low</b>	1.00	2.00	.27
	2.00	1.92	.61
	3.00	2.00	.68
	<b>Total</b>	<b>1.97</b>	<b>.54</b>
<b>Total 3 tasks</b>		<b>2.32</b>	<b>.65</b>

In Table 4.9, the high proficiency group has higher total grammar means than the low proficiency group ( $\bar{x}_H=2.66$ ,  $\bar{x}_L=1.97$ ), showing that the high proficiency group had better grammar performances than the low proficiency group. For the standard deviations, both proficiency groups have similar total spreading of the scores (High SD=.57, Low SD=.54), indicating that the variations of the scores from the two proficiency groups were not much different. The following table presents the effects of proficiency levels and task types on language functions mean scores.

**Table 4.10: The effects of proficiency levels and task types on language functions scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	16.26(a)	5	3.25	10.6	0	
Intercept	714.18	1	714.18	2327.25	0	
Proficiency levels	13.16	1	13.16	42.90	.00*	0.26
Task types	2.98	2	1.49	4.86	.01*	0.06
Proficiency levels * Task types	0.11	2	0.05	0.18	0.82	
Error	34.98	114	0.30			
<b>Total</b>	<b>765.43</b>	<b>120</b>				
<b>Corrected Total</b>	<b>51.24</b>	<b>119</b>				

\* $p \leq .05$

The results from Table 4.10 show a statistical significance at .05 level on both the proficiency levels,  $F(1, 114) = 42.90$ ,  $p < .05$  and task types,  $F(2, 114) = 4.86$ ,  $p < .05$ , indicating the main effects of these variables on the language functions performances. The effect size of the first variable is .26 whereas that of the second one is only .06. In other words, the two proficiency groups' language functions performances were significantly different among the three task types. Due to the statistical significance of the three task types, a Scheffé post-hoc test was performed to investigate the differences among the mean scores in the three task types from the two proficiency groups.

**Table 4.11: Scheffé post- hoc test on language functions scores**

(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.33(*)	.12	.027	.03	.64
	3.00	.33(*)	.12	.031	.02	.63
2.00	1.00	-.33(*)	.123	.027	-.64	-.03
	3.00	-.01	.12	.999	-.31	.30
3.00	1.00	-.33(*)	.12	.031	-.63	-.02
	2.00	.01	.123	.999	-.30	.31

\* $p \leq .05$ 

The results from Table 4.11 show that the significant difference in the mean scores between Task type one and two, and between Task type one and three is .33 ( $\bar{x}_{T1}=2.66$ ,  $\bar{x}_{T2}=2.33$ ,  $\bar{x}_{T3}=2.33$ ). The p values of these tasks are significant at .05 level. The means and standard deviations of the language functions scores are given below.

**Table 4.12: Descriptive statistics of language functions scores**

Proficiency levels	Task types	Mean	SD
<b>High</b>	1.00	2.95	.33
	2.00	2.67	.33
	3.00	2.68	.81
	<b>Total</b>	<b>2.77</b>	<b>.55</b>
<b>Low</b>	1.00	2.37	.38
	2.00	1.97	.57
	3.00	1.97	.70
	<b>Total</b>	<b>2.10</b>	<b>.59</b>
<b>Total</b>	1.00	2.66	.45
	2.00	2.33	.58
	3.00	2.33	.83
<b>Total 3 tasks</b>		<b>2.43</b>	<b>.66</b>

Table 4.12 demonstrates that the total of language functions mean scores of the high proficiency group is higher than that of the low proficiency group ( $\bar{x}_H=2.77$ ,  $\bar{x}_L=2.10$ ). This means that the high proficiency group performed better in language functions than the low proficiency group. The most difference in mean scores between the two proficiency groups is in Task type three ( $\bar{x}_H=2.68$ ,  $\bar{x}_L=1.97$ ) while the least difference is in Task type one ( $\bar{x}_H=2.95$ ,  $\bar{x}_L=2.37$ ). The low proficiency group poses the similar total of standard deviations to the high proficiency group (Low SD=.59, High SD=.55). The most difference in standard deviations is in Task type two (High SD=.33, Low SD=.57) whereas the least difference is in Task type one (High SD=.33, Low SD=.38). In other words, the two proficiency groups had the most different spreading of the scores in Task type two whereas the least different one was in Task type one. Table 4.13 illustrates the effects of proficiency levels and task types on cohesion mean scores.

**Table 4.13: The effects of proficiency levels and task types on cohesion scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	9.49(a)	5	1.89	6.76	0	0.22
Intercept	639.40	1	639.40	2279.67	0	
Proficiency levels	9.07	1	9.07	32.36	.00*	
Task types	0.40	2	0.20	0.72	0.49	
Proficiency levels * Task types	0.02	2	0.00	0.02	0.98	
Error	31.97	114	0.28			
<b>Total</b>	<b>680.87</b>	<b>120</b>				
<b>Corrected Total</b>	<b>41.46</b>	<b>119</b>				

\* $p \leq .05$

For the cohesion mean scores in Table 4.13, there is the only significant main effect at .05 level of the proficiency levels on the total cohesion scores,  $F(1,114) =$

32.36,  $p < .05$  with the effect size at .22. This means that the two proficiency groups' cohesion performances were significantly different. The mean scores and standard deviations of the cohesion scores are given below.

**Table 4.14: Descriptive statistics of cohesion scores**

Proficiency levels	Task types	Mean	SD
<b>High</b>	1.00	2.55	.41
	2.00	2.52	.43
	3.00	2.67	.75
	<b>Total</b>	<b>2.58</b>	<b>.55</b>
<b>Low</b>	1.00	2.02	.30
	2.00	1.97	.57
	3.00	2.10	.57
	<b>Total</b>	<b>2.03</b>	<b>.49</b>
<b>Total 3 tasks</b>		<b>2.30</b>	<b>.59</b>

Table 4.14 shows that the high proficiency group did better than the low proficiency group in cohesion from the total means of the two proficiency groups ( $\bar{x}_H=2.58$ ,  $\bar{x}_L=2.03$ ). In addition, their total standard deviations are not much different (High SD=.55, Low SD=.49), indicating that the variations of the scores from the two proficiency groups were not much different. Table 4.15 illustrates the effects of proficiency levels and task types on fluency mean scores.

**Table 4.15: The effects of proficiency levels and task types on fluency scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	14.83(a)	5	2.96	10.65	0	
Intercept	624.49	1	624.49	2242.93	0	
Proficiency levels	12.19	1	12.19	43.79	.00*	0.26
Task types	1.05	2	0.52	1.89	0.16	
Proficiency levels * Task types	1.58	2	0.79	2.84	0.06	
Error	31.74	114	0.27			
<b>Total</b>	<b>671.06</b>	<b>120</b>				
<b>Corrected Total</b>	<b>46.57</b>	<b>119</b>				

\* $p \leq .05$ 

Table 4.15 shows that the only statistical significance at .05 level is in the proficiency levels,  $F(1, 114) = 43.79$ ,  $p < .05$  with the small effect size at .26, showing the main effect of this variable on the fluency total scores. This indicates that the two proficiency groups' fluency performances significantly differed. The means and standard deviations of the fluency scores are given below.

**Table 4.16: Descriptive statistics of fluency scores**

<b>Proficiency levels</b>	<b>Task types</b>	<b>Mean</b>	<b>SD</b>
<b>High</b>	1.00	2.71	.46
	2.00	2.43	.45
	3.00	2.65	.70
	<b>Total</b>	<b>2.60</b>	<b>.55</b>
<b>Low</b>	1.00	1.75	.32
	2.00	1.96	.58
	3.00	2.17	.54
	<b>Total</b>	<b>1.96</b>	<b>.52</b>
<b>Total 3 tasks</b>		<b>2.28</b>	<b>.62</b>

Table 4.16 shows that the high proficiency group performed better than the low proficiency group in fluency ( $\bar{x}_H=2.60$ ,  $\bar{x}_L=1.96$ ). Additionally, the total standard deviations from the two proficiency groups are not much different (High SD=.55, Low SD=.52), showing that the variations of the scores from the two proficiency groups were not much different. The following table presents the effects of proficiency levels and task types on content knowledge mean scores.



**Table 4.17: The effects of proficiency levels and task types on content knowledge scores**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
Corrected Model	31.02(a)	5	6.20	25.34	0	
Intercept	553.62	1	553.62	2261.36	0	
Proficiency levels	18.60	1	18.60	75.99	.00*	0.32
Task types	12.27	2	6.13	25.05	.00*	0.21
Proficiency levels * Task types	0.15	2	0.07	0.31	0.73	
Error	27.90	114	0.24			
<b>Total</b>	<b>612.56</b>	<b>120</b>				
<b>Corrected Total</b>	<b>58.93</b>	<b>119</b>				

\* $p \leq .05$ 

For the last component, content knowledge, Table 4.17 shows that there are significant main effects at .05 level from both the proficiency levels,  $F(1,114) = 75.99, p < .05$  and task types,  $F(2,114) = 25.05, p < .05$  on the total content knowledge scores. The effect size of these variables is considerably small at .32 for proficiency levels and at .21 for task types. In other words, the content knowledge performances from the two proficiency groups significantly differed among the three task types. A Scheffé post-hoc test was conducted to investigate the mean differences among the three task types. The following table yields the results of the post-hoc test.

**Table 4.18: Scheffé post- hoc test on content knowledge scores**

(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	.67 (*)	.11	.00	.39	.94
	3.00	.69 (*)	.11	.00	.41	.96
2.00	1.00	-.67 (*)	.11	.00	-.94	-.39
	3.00	.02	.11	.98	-.25	.29
3.00	1.00	-.69 (*)	.11	.00	-.96	-.41
	2.00	-.02	.11	.98	-.29	.25

\* $p \leq .05$ 

Table 4.18 displays significant mean differences of content knowledge performed by the two proficiency groups between Task type one and two at .67, and Task type one and three at .69, ( $\bar{x}_{T1}=2.60$ ,  $\bar{x}_{T2}=1.93$ ,  $\bar{x}_{T3}=1.91$ ). All the p values are significant at .05 level. The mean scores and standard deviations of the content knowledge are given below.

**Table 4.19: Descriptive statistics of content knowledge scores**

Proficiency levels	Task types	Mean	SD
<b>High</b>	1.00	2.95	.34
	2.00	2.32	.43
	3.00	2.35	.72
	<b>Total</b>	<b>2.54</b>	<b>.59</b>
<b>Low</b>	1.00	2.25	.40
	2.00	1.53	.59
	3.00	1.47	.34
	<b>Total</b>	<b>1.75</b>	<b>.57</b>
<b>Total</b>	1.00	2.60	.51
	2.00	1.93	.65
	3.00	1.91	.71
<b>Total 3 tasks</b>		<b>2.14</b>	<b>.70</b>

From Table 4.19 the high proficiency group poses higher total of content knowledge means than that of the low proficiency group ( $\bar{x}_H=2.54$ ,  $\bar{x}_L=1.75$ ), indicating that the high proficiency group had better content knowledge performances than the low proficiency group. The most difference in the mean scores between the two proficiency groups is in Task type three ( $\bar{x}_H=2.35$ ,  $\bar{x}_L=1.47$ ) while the least difference is in Task type one ( $\bar{x}_H=2.95$ ,  $\bar{x}_L=2.25$ ). This means that the two proficiency group performed most differently in Task type three, while there were almost similar performances in Task type one. As for the standard deviations, the high proficiency group have similar total of standard deviations to that of the low proficiency group (High SD=.59, Low SD=.57), showing that they had similar spreading of the scores. The most difference in the spreading of the scores is in Task type three (High SD=.72, Low SD=.34) while the least difference is in Task type one (High SD=.34, Low SD=.40). The following part displays the mean difference of the LSP individual component in each proficiency group.

#### **4.1.3 The effect of task types on the LSP individual component in each proficiency level**

From the statistical significance at .05 level on the proficiency levels, the one-way ANOVA was conducted to investigate the effect of task types on the individual component in each proficiency group's performances. The results are presented in the following tables, and the mean difference in each LSP speaking component of the high proficiency students is illustrated in Table 4.20.

**Table 4.20: The summary table of the mean difference in each LSP speaking component of the high proficiency students**

Components		Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
<b>Pronunciation</b>	Between Groups	.38	2	.19	.71	.50	
	Within Groups	15.28	57	.27			
	<b>Total</b>	<b>15.66</b>	<b>59</b>				
<b>Vocabulary</b>	Between Groups	1.94	2	.97	3.32	.04*	0.10
	Within Groups	16.64	57	.29			
	<b>Total</b>	<b>18.59</b>	<b>59</b>				
<b>Grammar</b>	Between Groups	1.81	2	.90	2.94	.06	
	Within Groups	17.52	57	.31			
	<b>Total</b>	<b>19.33</b>	<b>59</b>				
<b>Lang Func</b>	Between Groups	.97	2	.48	1.65	.20	
	Within Groups	16.70	57	.29			
	<b>Total</b>	<b>17.67</b>	<b>59</b>				
<b>Cohesion</b>	Between Groups	.26	2	.13	.42	.66	
	Within Groups	17.70	57	.31			
	<b>Total</b>	<b>17.96</b>	<b>59</b>				
<b>Fluency</b>	Between Groups	.84	2	.42	1.35	.27	
	Within Groups	17.57	57	.31			
	<b>Total</b>	<b>18.40</b>	<b>59</b>				
<b>Cont Knowl</b>	Between Groups	5.01	2	2.50	8.98	.00*	0.24
	Within Groups	15.89	57	.28			
	<b>Total</b>	<b>20.90</b>	<b>59</b>				
<b>Total</b>	<b>Between Groups</b>	<b>.87</b>	<b>2</b>	<b>.43</b>	<b>1.77</b>	<b>.18</b>	
	<b>Within Groups</b>	<b>13.97</b>	<b>57</b>	<b>.25</b>			
	<b>Total</b>	<b>14.84</b>	<b>59</b>				

\* $p \leq .05$

Table 4.20 shows significant difference at .05 level between groups in vocabulary,  $F(2, 57) = 3.32, p < .05$  and content knowledge,  $F(2, 57) = 8.98, p < .05$ . This indicates the main effect of the task types on these two components. The effect size of vocabulary is .10 whereas the higher value is in content knowledge at .24. In other words, high proficiency students performed differently in vocabulary and content knowledge among the three task types. Due to the significant difference in mean scores, a Scheffé post-hoc test was performed to find the differences. Since the group was assigned by the task type, the means of the three task types were compared. The results of the mean comparison are illustrated in Table 4.21.

**Table 4.21: Scheffé post-hoc test in high proficiency groups**

Components	(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Vocabulary	1.00	2.00	.44(*)	.17	.04	.00	.86
		3.00	.26	.17	.31	-.16	.69
	2.00	1.00	-.44 (*)	.17	.04	-.86	-.00
		3.00	-.18	.17	.59	-.60	.25
	3.00	1.00	-.26	.17	.31	-.69	.16
Cont Knowl	1.00	2.00	.62(*)	.16	.00	.20	1.04
		3.00	.60(*)	.16	.00	.18	1.01
	2.00	1.00	-.62(*)	.16	.00	-1.04	-.20
		3.00	-.02	.16	.98	-.44	.39
	3.00	1.00	-.60(*)	.16	.00	-1.01	-.18
		2.00	.02	.16	.98	-.39	.44

\* $p \leq .05$

In Table 4.21, the results of Scheffé post-hoc test demonstrate that the vocabulary mean scores in Task type one differ from Task type two at .44 ( $\bar{x}_{VT1}=2.90, \bar{x}_{VT2}=2.46$ ), indicating the different vocabulary performances between these task types. For the content knowledge, the means in Task type one differ from Task type two at .62 and from Task type three at .60 ( $\bar{x}_{CNT1}=2.95, \bar{x}_{CNT2}=2.33, \bar{x}_{CNT3}=2.35$ ). All the p values are significant at .05 level. In other words, the high proficiency group had different content knowledge performances between these task types. The following

table presents the mean difference in each speaking component of the low proficiency students.

**Table 4.22: The summary table of the mean difference in each LSP speaking component of the low proficiency students**

Components		Sum of Squares	df	Mean Square	F	Sig.	$\eta^2$
<b>Pronunciation</b>	Between Groups	.20	2	.10	.32	.72	
	Within Groups	18.65	57	.32			
	<b>Total</b>	<b>18.85</b>	<b>59</b>				
<b>Vocabulary</b>	Between Groups	1.40	2	.70	2.08	.13	
	Within Groups	19.27	57	.33			
	<b>Total</b>	<b>20.68</b>	<b>59</b>				
<b>Grammar</b>	Between Groups	.07	2	.03	.12	.88	
	Within Groups	17.51	57	.30			
	<b>Total</b>	<b>17.58</b>	<b>59</b>				
<b>Lang Func</b>	Between Groups	2.13	2	1.06	3.33	.04*	0.10
	Within Groups	18.28	57	.32			
	<b>Total</b>	<b>20.42</b>	<b>59</b>				
<b>Cohesion</b>	Between Groups	.15	2	.07	.32	.73	
	Within Groups	14.27	57	.25			
	<b>Total</b>	<b>14.43</b>	<b>59</b>				
<b>Fluency</b>	Between Groups	1.80	2	.90	3.63	.03*	0.11
	Within Groups	14.17	57	.24			
	<b>Total</b>	<b>15.97</b>	<b>59</b>				
<b>Cont Knowl</b>	Between Groups	7.41	2	3.70	17.58	.00*	0.38
	Within Groups	12.02	57	.21			
	<b>Total</b>	<b>19.43</b>	<b>59</b>				

<b>Total</b>	<b>Between Groups</b>	<b>.37</b>	<b>2</b>	<b>.18</b>	<b>.84</b>	<b>.44</b>
	<b>Within Groups</b>	<b>12.73</b>	<b>57</b>	<b>.22</b>		
	<b>Total</b>	<b>13.10</b>	<b>59</b>			

\* $p \leq .05$

Table 4.22 shows the significant difference at .05 level between groups in the low proficiency students in the language functions,  $F(2, 57) = 3.33, p < .05$ , fluency,  $F(2, 57) = 3.63, p < .05$  and content knowledge,  $F(2, 57) = 17.58, p < .05$ , indicating the task types effect on three language components. The effect size of the first two components is not much different ( $\eta^2_{LF} = .10$  and  $\eta^2_F = .11$ ) while the highest value is in content knowledge ( $\eta^2_{CK} = .38$ ) which is similar to the high proficiency group. From these statistical results, the performances of the low proficiency students in these components significantly differed among the three task types. A Scheffé post-hoc test was conducted to find the differences among the means from the three groups. The results of the analysis are provided below.

**Table 4.23: Scheffé post-hoc test in low proficiency groups**

Components	(I) task	(J) task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
<b>Lang Func</b>	1.00	2.00	.40 (*)	.17	.02	.04	.75
		3.00	.40(*)	.17	.02	.04	.75
	2.00	1.00	-.40 (*)	.17	.02	-.75	-.04
		3.00	.00	.17	1.00	-.35	.35
		3.00	1.00	-.40 (*)	.17	.02	-.75
	2.00	.00	.17	1.00	-.35	.35	
<b>Fluency</b>	1.00	2.00	-.21	.15	.40	-.60	.18
		3.00	-.42(*)	.15	.03	-.82	-.02
	2.00	1.00	.21	.15	.40	-.18	.60
		3.00	-.21	.15	.40	-.60	.18
		3.00	1.00	.42(*)	.15	.03	.02
	2.00	.21	.15	.40	-.18	.60	
<b>Cont Knowl</b>	1.00	2.00	.71 (*)	.14	.00	.34	1.07
		3.00	.77(*)	.14	.00	.41	1.14
	2.00	1.00	-.71 (*)	.14	.00	-1.07	-.34
		3.00	.06	.14	.91	-.30	.42
		3.00	1.00	-.77(*)	.14	.00	-1.14
	2.00	-.06	.14	.91	-.42	.30	

\* $p \leq .05$ 

Table 4.23 demonstrates that the mean difference of language functions between Task type one and two is .40. The same value occurs between Task type one and three. The means of the three tasks are 2.37, 1.97 and 1.97 respectively. For fluency, the difference in the means between Task type one and three is .42 ( $\bar{x}_{FT1}=1.75$ ,  $\bar{x}_{FT3}=2.17$ ). For content knowledge, the mean difference between Task type one and two is .71, and Task type one and three is .77 ( $\bar{x}_{CNT1}=2.25$ ,  $\bar{x}_{CNT2}=1.54$ ,  $\bar{x}_{CNT2}=1.48$ ). All the p values are significant at .05 level. The following part displays the content analysis of the speaking performances from the two proficiency groups.



#### **4.1.4 Content analysis of the LSP speaking performances from high and low proficiency groups**

From the significant differences between the total LSP speaking performances of the high and low proficiency groups, content analysis was conducted to investigate both the similarities and differences in each LSP speaking component of the test performances. The analysis also showed the in-depth information and the prominent features in some of the LSP speaking components associated with the proficiency group and a particular task type. Additionally, typical errors and the errors that were prominent in a particular task type and the proficiency group were found. The following identification of the students was used in the responses. The letter 'A' represents 'High proficiency group' while the letter 'B' indicates 'Low proficiency group'. The number following the letter represents the task and the last number indicates the order of the individual student. The information in the brackets was added for the clarity of the responses.

- **Pronunciation**

Pronunciation is the first linguistic component of the language knowledge. The investigation is on accuracy of pronouncing words and the accurate use of stress and intonation in the speech. The two proficiency groups made similarly typical errors in the incorrect pronunciation of whole words, endings of words and wrong stress in their responses across the three task types. However, the salient difference in pronunciation between the two proficiency groups was in errors that were mainly found in the low proficiency students' responses across the three task types. The examples of these errors are presented in relation to each task type as follows.

In Task type one, the most prominent errors occurred in the wrong pronunciation of the whole words and mispronouncing the endings of the words in the students' responses as illustrated in the following excerpts from the two proficiency groups.

Excerpts	Incorrect pronoun. words	Incorrect pronouncing endings	Stress
		/-z/, /-st/, /-t/, /-dʒ/, /-ʃ/,	
<b>Student A1.4</b>			
<p>Picture1: This is the Emerald Buddha Temple. It <i>wat</i>[was] built in the reign of King Rama the <i>firt</i> [first] in 1782. The temple is very important because it is believed to be the most sacred place[s] in Thailand. There are many interesting [things] to see inside for <i>ikxample</i>[example] <i>the</i>[/ðɪ:/] Emerald Buddha <i>imade</i>[image], the ordination hall and the gallery.</p>	✓	✓ /-z/, ✓ /-st/	
<p>Picture 2: This is <i>the</i>[/ðɪ:/] emerald Buddha image. It <i>wat</i> [was] found in 1434. It <i>wat</i> [was] made from green jade. It is in [pause] meditation posture. It is 48.3 <i>centimet</i>[ters] wide and high 66 <i>centimet</i>[ters]. It has 2 seasonal costumes made in the reign of King Rama the 1<sup>st</sup>, one for summer and one for the rainy season. It is now <i>enslide</i> [enshrined] inside the ordination hall.</p>	✓	✓ /-z/	
<p>Picture4: There are 178 the mural painting[s] at the gallery. There [They] are located opposite to <i>the</i>[/ðɪ:/] ordination hall. They were painted in the reign of King Rama the first and <i>lanovated</i>[renovated] many time [s]. The story of the painting was about the Rama Yana or Ramakien.</p>	✓		
<p>Picture5: This is the Grand Palace. It was built in [the] reign of King the 1<sup>st</sup> in 1782. The Grand Palace is very important because it is [the] residence of King Rama.</p>	✓		
<p>Picture 6: ... It is very important because there is พระที่นั่งจักรีมหาปราสาท which is one of the most sacred[places] in Thailand and Thai people pay respect... There are many interesting [things] to see inside such as the top floor of the central mansion are <i>krep</i> [kept] in the royal <i>ad</i> [ash]. [At] the top floor of the eastern wing <i>reallycios</i> [religious] objects are <i>krep</i> [kept].</p>	✓	✓ /-ʃ/	
<p>Task 2</p> <p>At 8 o'clock depart from The Grand Hotel Bangkok. Next at 930 arrive we will arrive at Nakhon Pathom visit golden Pagoda and pay respect to the sacred Buddha <i>imade</i> [image]. Then, at[pause] 1030 we will Visit[pause] Dvaravati Museum. [pause] After that at 1130 o'clock sightseeing and we will [go] sightseeing</p>	✓	✓ /-dʒ/	

<p>and buy a <i>sound venic</i> [souvenirs] at the Road at the Local market. Then 1245 o'clock we will have lunch at 'Ban Ruen Thai Restaurant', a famous restaurant in Nakhon Pathom, that offers the delicious local dishes such as grilled river[<i>pause</i>] <i>pround</i> [prawn], spicy salad and <i>firt</i> and <i>firt</i> [fresh] [<i>pause</i>] <i>skilt</i> [squid], fried <i>children</i> [chicken] with Thai herbs, and <i>spitch</i> [spicy] [<i>pause</i>] Nakhon Pathom soup. Next, 1345 we will visit[<i>pause</i>] <i>Sa Sa Sanoom Chandra</i> [Sanamchandra] <i>Palate</i>[Palace]'....</p>	<p>✓  ✓ ✓ ✓ ✓ ✓ ✓</p>	<p>✓ /-st/</p>	
<p><b>Student B1.1</b></p> <p>Picture1: This is the Emerald Buddha Temple. It was built in the reign of <i>Kring</i> [King] Rama [the] first. It <i>wat</i> [was] built by King Rama the first. It is the symbol of Thai nation. There are many interesting [things] to see; for example, the ordination [<i>pause</i>] hall and the Emerald Buddha <i>imesh</i> [image].</p> <p>Picture 2: Now the emerald Buddha imesh [image]. It <i>wat</i> [was] made from jade. It is in the <i>postore</i> [posture] of meditation. It is 48.3 cms wide and 66 cms tall. It was found in Chiang Rai in the norther[n] [of] Thailand. It is enshrined in the ordination hall.</p> <p>Picture3: ... There are many interesting [things] to see for example the <i>best</i> [base] of the ordination hall and garudas holding naga. Several <i>mewrals</i> [murals] and the gallery outside.</p> <p>Picture4: This is the gallery. It was built in the reign of <i>Kring</i> [King] Rama [the] first. It <i>wat</i> [was] built by King Rama the first. It is note[ed] for its murals depicting the entire รามเกียรติ์. There are many interesting [things] to see for example the inside wall with <i>mewrals</i> [murals]...</p> <p>Picture5: This is the Grand Palace. It was built in 1782. It was built by King Rama the 1<sup>st</sup>. It is a <i>residedence</i>[residence] of King Mongkut Rama the third until the entire age of his life. There are many interesting [things] to see for example <i>the</i>[/ði:/] พระมหามณเฑียร group, พระมหาปราสาท group, พระที่นั่งจักรีมหาปราสาท group, บรมพิมาน Mansion and สีวาลัย group.</p> <p>Picture 6: This is the พระที่นั่งจักรีมหาปราสาท. It was built in the reign of King จุฬาลงกรณ์. It was built by King Rama the fifth. It is a reception hall for royal <i>great</i> [guest]. There are many interesting [things] to see for example in the rear center of จักรีมหาปราสาท is the Chakri Thorn [Throne] room and the</p>	<p>✓    ✓  ✓  ✓  ✓</p>	<p>✓ /-z/  ✓ /-dʒ/  ✓ /-z/  ✓ /-st/  ✓ /-z/</p>	<p>✓</p>

symbol of Chakri <i>dynasty</i> [dynasty] is present[ed] on the wall behind the thorn [throne].	✓		
Picture7: This is the Dusit Mahaprasat <i>Thorn</i> [Throne] Hall. It was built in seven [pause] 17 [pause] 89. It <i>wat</i> [was] by King Rama the 1 <sup>st</sup> . It [is] used for annual <i>consignation</i> [consecration] day ceremony. ...	✓ ✓	✓ /-z/	
Task 2:  Bangkok-Nakhon Pathom .We will depart from The Grand Hotel Bangkok at 08 o'clock. Next, we will arrive at Nakhon Pathom visit <i>growden</i> [golden] Pagoda and <i>pray</i> [pay] [pause] respect to the[pause] <i>re-cord</i> [record]Buddha <i>imesh</i> [image].  ... <i>Lart</i> , [Last] [Go] <i>siteseeting</i> [sightseeing] and buy souvenirs at the Local market at [pause]...	✓ ✓	✓ /-dʒ/, ✓ /-st/, ✓ /-t/	✓

(The italic words represent errors in pronunciation. Grammatical errors are not corrected. Only the errors related to the topic are presented in the italic form.)

In Task type one, most of the errors were in the incorrect pronunciation of words and endings. As for the error in incorrect pronouncing words, most of the errors were found in the technical terms; for example, the words ‘renovate’, ‘centimeter’ and ‘religious’ from the student A1.4; and ‘posture’, ‘murals’, ‘dynasty’ and ‘consecration’ from the student B1.1. These words should have been correctly pronounced as /'ren.ə.veɪt/, /sentɪ'mi:tə/, /rɪ'laɪdʒ.əs/, /'pɑ:s.tʃə/, /'mjʊr.əl/, /'daɪ.nəsti/ and /kɑ:nt.sɪ'kreɪ.ʃən/. For the incorrect pronunciation of endings, student A1.4 mispronounced the sounds /dʒ/, /-st/, /-z/ and /-ʃ/ on the words ‘image’, ‘first’, ‘was’ and ‘ash’. The correct pronunciation should be /'ɪm.ɪdʒ/, /'fɜ:st/, /wɑ:z/ and /æʃ/. The student B1.1 wrongly pronounced the sounds /-z/, /-dʒ/, and /-st/ in the words ‘was’, ‘image’, and ‘last’. The correct pronunciation should be /wɑ:z/, /'ɪm.ɪdʒ/, and /læst/. All these selected words were frequently mispronounced in this task type. The error in stress was only found in the low proficiency student B1.1 as in the words ‘residence’ and ‘record’ that were wrongly stressed as /re'z.ɪ.dənts/ and /'rɪkɔ:rd/ instead of /'rez.ɪ.dənts/ and /rɪ'kɔ:rd/.

In Task type two, errors in the pronunciation of words, endings, and stress were also found. Some of them are given below.

Excerpts	Incorrect pronouncing words	Incorrect pronouncing endings /-s/,/-z/,/-t/, /-d/,/-ʃ/, /-ŋ/,/dʒ/, /-l/	Stress
<p><b>Student A2.5</b></p> <p>Please do not short <i>prant</i> [pants] and a spaghetti shirt because it is impolite in Thai culture.</p> <p>Please do not <i>climbing</i>[/klaɪmbɪŋ/] Buddha image because it is impolite in Thai culture and dangerous.</p> <p>Please do not sit the impolite posture in front of Buddha <i>satstus</i> [statue]because it is bad in Thai culture.</p> <p>Please be call police if you any problem stolen wallet lost <i>variable</i> [valuable] belongings because police can help you.</p> <p>Please be you have hotel <i>gard</i> [card] because you can going anywhere.</p> <p>Please don't <i>brink</i> [bring] belonging inside the <i>bud</i> [bus] because it is rules.</p> <p><b>Student A2.6</b></p> <p>Any <i>problem</i> [/'pra:.bləm/] stolen wallet lost <i>vorible</i> [valuable] belonging.</p>	<p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p>	<p>✓ /-ŋ/, ✓ /-s/</p>	<p>✓</p>
<p><b>Student B2.1</b></p> <p>Any problem[s] <i>stole</i>[stolen] wallet, lost <i>value ble</i> [valuable] belonging and [pause] and call me the tourist police.</p> <p><b>Student B2.4</b></p> <p>You should not klimbing[/klaɪmbɪŋ/] on the Buddha <i>imade</i>[image] because Thai people respect to the Buddha <i>imade</i>[image].</p> <p>You should carry the map because you don't get <i>lot</i>[lost].</p>	<p>✓</p> <p>✓</p>	<p>✓ /dʒ/ ✓ /dʒ/  ✓ /-st/</p>	<p>✓</p>

You should not put your belonging on the <i>bud</i> [bus] because it safety and your belonging.		✓ /-s/	
<b>Student B2.6</b>			
Please do not take please do not take a photo <i>becaud</i> [because] it is the <i>ruse</i> [rule].		✓ /-z/, /-l/	
Do not stop mouth [pause] <i>becaud</i> [because] polite <i>becaud</i> [because] it is polite.		✓ /-z/	
Please do not walk on the <i>yars</i> [yard][pause] <i>becaud</i> [because] it is the <i>ruse</i> [rules].		✓ /-d/, /-z/, ✓ /-l/	

(The italic words represent errors in pronunciation. Grammatical errors are not corrected. Only the errors related to the topic are presented in the italic form.)

In Task type two, the majority of the students made the error in the incorrect pronunciation of endings, and the examples were from the student B2.6. This one incorrectly pronounced the ending sound /-z/ in the word ‘because’ as /bikəd/ instead of /bikəz/ and ‘rules’ as /ru:z/ instead of /ru:lz/. Similarly, the student B2.4 mispronounced the ending sound of /- dʒ/ in the word ‘image’ as /ɪm.ɪd/ instead of /ɪm.ɪdʒ/. The sound /-st/ was also mispronounced in the word ‘lost’ as /lɑ:t/ instead of /lɑ:st/ and the word ‘bus’ was incorrectly pronounced as /bʌd/ instead of /bʌs/. The ending /- ŋ/ in the word ‘bring’ was also mispronounced by the majority of the students in the word ‘bring’ as /brɪnk/ instead of /brɪŋ/ as in the student A2.5’s excerpt. The analysis also showed the error in the incorrect pronouncing words and the examples were from this student. The word ‘pants’ was incorrectly pronounced as /prænts/ instead of /pænts/, ‘statue’ as /sætstʌs/ instead of /stætʃ.u:/ and ‘valuable’ as /ve.ri.ə.bɪ/ instead of /væl.jʊ.bɪ/. The words ‘card’ and ‘climb’ were the two most mispronounced words by almost all of the students. The first one was mispronounced as /gɑ:rd/ instead of /kɑ:rd/ and the second one as /ˈklɪm.ɪŋ/ instead of /ˈklaɪ.mɪŋ/. For the stress error, the word ‘problem’ was incorrectly stressed on the second syllable as /prɑ:ˈbləm/ instead of /ˈprɑ:bləm/ by the student A2.6. Similarly, the student B2.1 put the wrong stress on the last syllable in the word ‘valuable’ as

/væl.jʊ.'bl/ instead of /'væl.jʊ.b|/. Errors in pronouncing words, endings and stress were also found in Task type three and some of them are provided below.

Excerpts	Incorrect pronouncing words	Incorrect pronouncing endings /-d/, /-g/, /-ndʒ /	Stress
<p><b>Student B3.3</b></p> <p>Oh sorry is my credit <i>gard</i>[card] use for buy for buy shirt is now I take going to police station for for talking about credit <i>gard</i>[card] you OK?</p> <p>I am so sorry. It don't <i>happle</i>[happen] again.</p> <p>Thank you. OK no <i>prompram</i>[problem]. I change with your room OK?</p> <p><b>Student B3.5</b></p> <p>I so sorry sir and madam. I will <i>shenk</i>[change] a new bus to transfer <i>light</i>[right] now. I will never happen again.</p> <p><b>Student B3.8</b></p> <p>Ok madam I will check <i>back</i>[bag] for you</p> <p><b>Student B3.15</b></p> <p>Ok madam I will call the police for you. Please tell me where did you forget your <i>wa_let</i> [wallet]?</p>	<p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p>	<p>✓ /-ndʒ /</p> <p>✓ /-g /</p>	<p>✓</p>

(The italic words represent errors in pronunciation. Grammatical errors are not corrected. Only the errors related to the topic are presented in the italic form.)

In Task type three, the error in incorrect pronunciation of words was prominent, the student B3.3 mispronounced a number of words such as pronouncing 'card' as /gɑ:rd/ instead of /kɑ:rd/, 'happen' as /hæp.ɪ/ instead of /hæp.ən/ and 'problem' as /prɔ:mpræm/ instead of /pra:bləm/. The student B3.5 mispronounced the word 'right' as /laɪt/ instead of /raɪt/ and the sound /-ndʒ/ in the word 'change' that was mispronounced as /tʃeɪnk/ instead of /tʃeɪndʒ/. The student B3.8 mispronounced the sound /-g/ in the word 'bag' as /bæk/ instead of /bæg/. The stress

error was found in the word ‘wallet’ which was incorrectly stressed on the second syllable as /wɑ:.'lɪt/ instead of /'wɑ:.'lɪt/ by the student B3.15.

Concerning the errors that were noted in a particular task type, there was no intonation in the responses which caused monotonous and unnatural speeches. The errors in incorrect pronunciation of Thai words and incorrect pronunciation of consonant clusters were also found, particularly in the mispronunciation of Thai words that were mainly made by the low proficiency students in Task type one. Examples are given below.

Students	Excerpts	Types of errors		
		No intonation	Incorrect pronoun. Thai words	Incorrect pronoun. consonant clusters
B1.4	Picture5: This is the <i>Gand</i> [Grand] Pace [Palace]. Picture6: ...the top <i>for</i> [floor] of the cen centast [central] mansion are crep [kept] the royal as [ashes] the top <i>for</i> [floor] of the[di] eastern wink [wing] object[s] are crep [kept] on the top <i>for</i> [floor] of western wink [wing] as [ashes] are crep [kept] of the royal queen.	✓		✓ /gr-/ ✓ /fl-/ ✓ /fl-/ ✓ /fl-/
B1.2	Picture7: Picture7: We are now standing in front of the Dusit Mahaprasat <i>Tone</i> [Throne] Hall...			✓ /θr-/
B1.3	Picture5: Next, we are going to see the Grand Prace [Palace]. It's near the Imerald[Emerald]Buddha [pause]Temple. It began to construct in [pause] seventeen [pause] eighty two and wash [was] <i>competed</i> [completed] in the rest [reign] of King Rama the first.  Picture3: This id [is] the ordination hall. [pause] It id [is] <i>contructed</i> [constructed] in 1782 by King Rama the firt [first].	✓		✓ /pl-/ ✓ /str-/



B1.2	Task2: Third 1030 o'clock we will visit <i>Davara Davara</i> [/dʌ? wā: rá? wá? tī:/] Museum. ...Then [pause] eleven forty five o'clock [1345] we will visit ' <i>Sanam Chada</i> ' [/sà? nə:m cān drā:/]		✓	
B1.5	Picture3: There are many interesthink [interesting] things to see inside such as ' <i>Phrasamputtapani</i> ' [/phá? phút thá? sām pān nī:i/].		✓	
A2.6	You should take off your shoes because [pause] we should respect this place and if you wear shoes <i>for</i> [floors] are dirty.			✓ /fl-/

(The italic words represent errors in pronunciation. Grammatical errors that do not cause misunderstanding are not corrected. Only the errors related to the topic are presented in the italic form.)

The first error in no intonation was noted in Task type one in the low proficiency responses by the students B1.3 and B1.4. They produced a monotonous speech and did not show any evidence on using the intonation. This error obstructed the natural flow of the speech and caused difficulty in understanding the speech.

The second and most prominent error was the incorrect pronunciation of consonant clusters in Task type one and two. The majority of the students did not pronounce the sound /l/ and /r/ with the previous consonants. The most prominent incorrect pronunciation was the sound /fl-/. The student B1.4 pronounced the word 'floor' as /fɔ:r/ instead of /flɔ:r/ which was similar to the student A2.6. The sound /θr-/ was also incorrectly pronounced by the student B1.4 in the word 'throne' as /təʊn/ instead of /θrəʊn/. The sound /gr-/ was also incorrectly pronounced by the student B1.4 as /gænd/ instead of /grænd/. The error was found in the /pl-/ sound in B1.2's excerpt that the word 'complete' was incorrectly pronounced as /kəmpi:t/ instead of /kə'mpli:t/. The error in the three syllables consonant cluster was found in the sound of /str-/ in the student B1.3's excerpt. This student left both

/s-/ and /r-/ sounds from the word ‘construct’ as /kə'ntʌkt/ instead of /kə'nstrʌkt/. The last error was the incorrect pronouncing Thai words which occurred only in Task type one in both the high and low proficiency students. The majority of the students incorrectly pronounced the word ‘Davaratavati’. The student B1.2 pronounced part of the words as /dā: wā: rā:/ instead of /dā: wā: rá? wá? tī:/ or /thá? wā: rá? wá? dī:/. Another frequent mispronunciation was the word ‘Sanamchandra’ that can be seen from the same student B1.2 as /sà? nǎ:m cā dā:/ instead of /sà? nǎ:m cān drā:/ or /sà? nǎ:m cān/. The last example was the word ‘Phraputtasampanni’ which was one of the most mispronounced words from both the high and low proficiency groups. The student B1.5 pronounced this word as /phá? sām phút thá? phā: nī:/ instead of /phá? phút thá? sām pān nī:/. The error in incorrect pronunciation of Thai words in Task type one may be from the types of test contents and proficiency levels in the target language. This error was mainly made by the low proficiency students. It was also clearly seen that only Task type one required the students to explain about the attractions, particularly temples and palaces, and these places were in Thai names.

- **Vocabulary**

The investigated features in the use of vocabulary were technical and generic terms in the speech. It was measured by the accuracy and range in the responses. Range of vocabulary was measured in number of words per response. The analysis showed that the two proficiency groups employed similar types of vocabulary in their responses across the three task types. However, the salient difference in the vocabulary performances between the two proficiency groups was in range and accuracy. The high proficiency group used a wider range of vocabulary than their counterparts whereas most of the errors were mainly found in the low proficiency students’ responses.

Regarding the types of vocabulary, the two proficiency groups used similar tourism-related technical terms, particularly about Thai history and architectural structures in Task type one and some of them were limitedly used in Task type two.

The reasons may be that in Task type one, the students were required to explain about the attractions on Thai architecture, arts, history and Buddhism. As for Task type two, they were required to give suggestions to the tourists on do's and don'ts which were related to Thai etiquettes. This information was associated with technical terms. The generic terms were also used in Task type two. However, most of the generic terms were found in Task type three in which the students were asked to resolve the problems in the organized trip which was related to general information. The following excerpts demonstrate the technical terms in italic used by the two proficiency groups in Task type one.

A1.3: Now, I'd like to give you the brief history of the *Emerald Buddha image*. It is carved from a large piece of green jade. It was found in a *Buddha statue* in a *Chedi* in Chiang Rai in [the] northern [part] of Thailand. The image itself date[s] back more than six hundred years. It had been removed to many places both inside and outside the country such as Laos and Burma. Finally, it is *enshrined* in Thailand.

B1.4 Did [This] id [is] the *Emerade* [Emerald] *Buddha Temple*. Id [it] ward [was] built in the *rain* [reign] of *King Rama* the firt [first] in seventeen-eighty two. The *temple* id [is] very impression.

From the two excerpts, the high proficiency student (A1.3) used a wider range of vocabulary in the responses than the low proficiency student (B1.4), indicating the salient difference in the aspect of range. Additionally, another noted difference was in accuracy whereby all of the vocabulary errors were mainly found in the responses of the low proficiency students. These errors were the inaccurate use of words that caused incomprehensible, unclear and unintended meaning of the responses. Use of inaccurate words was from the similar pronunciation of the intended words and the near synonyms. The error also included using wrong words that were associated with the context. The following excerpts demonstrate the errors from the two proficiency groups in the three task types.

Students	Excerpts	Inaccurate use of words with similar pronunciation	Inaccurate use of words with near synonym	Inaccurate use of words with similar context
<b>B1.4</b>	<p>Picture 2: ..It's It has [pause] 2 seasonal costume[s] made [made] in the rain [reign] of King Rama the first [first] one for summer and one for [the] rainy season. It is now <i>excite</i> [exhibited] inside the[di] ordination hall.</p> <p>Picture4: ..[pause] the story [story] of the painting [painting] was about is Ramayana or Rama Ramakien the painting [s] are important because [because] they [pause] <i>replace</i> [reflect] Ratanakosin art.</p>	Excite for exhibited		
<b>A2.1</b>	<p>Please don't wear <i>spaghetti shirt</i> [singlet] and jean short[s] inside because it is impolite.</p> <p>Please go to follow <i>the map</i>[hotel card]because it will lead you the hotel</p>		Spaghetti shirt for singlet  Map for hotel card	
<b>A2.3</b>	Do not <i>include</i> [wear] the singlet and shorts because disrespectful			Include for wear
<b>A2.7</b>	<p>Please quiet because this place want [wants] <i>pacific</i> [peace]</p> <p>Please check <i>luggage</i>[bags, belongings] before you get off the bus</p>	Pacific for peace	Luggage for bags, belongings	
<b>A2.14</b>	If you [have] any problem please be <i>connect</i> [contact] to the tourist police number 1155 because we can take care and help you.	Connect for contact		
<b>A2.15</b>	You should beware pickpocket because this is <i>crown people</i> [crowded, crowded place]	Crown people for crowded		

<b>B2.1</b>	<p>Please do not wear <i>spaghetti shirt</i> [singlet] and shorts in the the tepen [temple]because it is noly [holy]</p> <p>Please do not [pause] the <i>photoshop</i>[take the photo] because it's [the] rule</p> <p>Please do not [pause] speak [pause] <i>power</i>[loudly] because it's [the] rule</p> <p>Please do not [pause] <i>walk</i> [climb] in the Buddha because it's [the] rule</p>		<p>Spaghetti shirt for singlet</p> <p>Photoshop for take the photo</p> <p>Power for loudly</p>	<p>Walk for climb</p>
<b>B2.3</b>	<p>Please always carry the <i>map</i>[hotel card] becaud [because] you will be get lost</p>		<p>Map for hotel card</p>	
<b>B2.18</b>	<p>Please do not [pause] <i>sound</i> [speak loudly] becaud [because] it is [the] rule</p>	<p>Sound for speak loudly</p>		
<b>B3.6</b>	<p>I will can[call] the police please where do you forget the <i>travel</i> [wallet]?</p>			<p>Travel for wallet</p>
<b>B3.17</b>	<p>I'm so sorry sir and madam. I will shenk[change] new <i>order</i> [room]. It will never happen again.</p>			<p>Order for room</p>

(The italic words represent errors in vocabulary. Grammatical errors are not corrected.)

From the data presented above, the majority of the students misused the vocabulary by replacing the similar pronunciation words with the correct words. They also used the near synonyms for the intended words. In addition, they misused the words that they could recall to respond to the test tasks and some of the words associated with the context; for example, the word 'walk' was used for 'climb' by the student B2.1 and 'travel' was for 'wallet' by the student B3.6.

Moreover, the prominent error in particular task types was in the use of generic terms that obstructed the clarity of the speech which was mainly found in Task type two.

A2.1: Please [take] *free service* to the JJ Mall near the tourist police office because tourist police in Thailand want [to] take care everybody.

B2.12: Please [take] *free service* to JJ Mall near the tourist police office.

A2.9: Please do not *disturb* in the ordination hall because we should pay respect this place.

B2.5: You should *behave* in the ordination hall because it's the sacred.

From the first two excerpts, the students A2.1 and B2.12 used the word 'free service' for 'transfer' to suggest the tourists for convenient travelling. However, the word 'free service' was too general and could be interpreted in various meanings other than the required information. For example, the free service could be interpreted from the context as the service from the tourist police which was incorrect; thus, it caused misunderstanding to the tourists. Another example was related to do's and don'ts at the religious site and the student A2.9 used the word 'disturb' and B2.5 used 'behave'. These two words could cause misunderstanding to the tourists since these words could be referred to a number of possible actions including dressing politely or sitting and standing politely. Although the meaning was not totally incorrect, it was not clear to the audiences or the tourists.

- **Grammar**

Grammar was investigated on the accuracy, range and complexity of the structures in the responses, particularly on the use of the tenses and types of sentences. Range of tenses and structures was measured in number of types of tenses and structures per response.

The analysis from the frequency counts of tenses in the responses of two proficiency groups showed that present simple and future tenses were mainly used across three task types. In Task type one, present simple tense was found at 825

frequencies, 386 frequencies in Task type two, and 230 frequencies in Task type three. For future tense, two proficiency groups used this tense at 283 frequencies in Task type one, 54 frequencies in Task type two and 285 frequencies in Task type three. However, past simple tense, particularly the passive voice was mainly found in Task type one in the responses of the two proficiency groups (Past simple tense=468 frequencies). For types of sentences, the analysis showed that simple and compound constructions were primarily found in Task type one and three whereas complex constructions were mainly used in Task type one and two. In Task type one, the two proficiency groups used 963 frequencies of simple constructions and 474 of them were found in Task type three. For compound constructions, 97 frequencies were found in Task type one and 35 of them were used by the two proficiency groups in Task type three. For complex constructions, the two proficiency groups used 148 frequencies of these sentences in Task type one and 395 of them were found in Task type two.

The use of a particular tense and sentence constructions may have stemmed from the information requirement in the task. In Task type one, students used present simple tense to explain details and the interesting things to see at attractions while both past simple and passive voice were used to explain the information that was related to history of the attraction such as year of construction as illustrated in the following excerpts. In Task type two, present simple tense was mainly used to provide polite suggestions to tourists while future 'will' was mainly found in Task type three to inform tourists that the students will take actions to resolve problems. The use of a particular type of sentences may stem from task requirement. A salient example was from Task type two in which complex sentences were mainly used by the two proficiency groups to provide the reason for their suggestions to the tourists. Moreover, the analysis revealed that the two proficiency groups used a particular structure in a particular task type which may be due to the information required in the task. In Task type one, the two proficiency groups similarly used particular structures to give the required information on the name of the place, year of construction and the tour program. In Task type two and three, the two proficiency groups used particular structures in relation to the task content requirement and it was noted that the high

proficiency group employed more sophisticated constructions (particularly the compound-complex sentences) than the low proficiency group in the three task types. In Task type one, the frequency counts showed that the high proficiency group employed 43 frequencies of compound-complex sentences whereas 20 of them were found in the low proficiency group's responses. In Task type two, the high proficiency group used 11 compound-complex sentences while three of them were employed by the low proficiency group. In Task type three, this construction was mainly used by the high proficiency group for four frequencies while none of this construction was found in the low proficiency group's responses. The examples on the use of tenses, types of sentences and particular structures from the two proficiency groups are presented below.

Excerpts	Grammatical features		
	Tenses	Types of sentences	Particular structures
<b>Task type one (High group):</b>			
This <b>is</b> the Emerald Buddha Temple.	Present simple	Simple	<sup>1</sup> Pronoun+ V to be + Article + Noun
It <b>was built</b> in the reign of King Rama the first in 1782.	Past simple	Simple	<sup>2</sup> Pronoun+ V to be + Past participle+ Complement
The Emerald Buddha Temple <b>was</b> very important because Thai people <b>believed</b> that it <b>was</b> the most sacred place[s] in Thailand and the repository of spirits for all Thai people.	Past simple	Compound-complex	<sup>3</sup> Noun+ V to be+ Quantifier+ +Adj+Conjunction+ Complement
There <b>are</b> many interesting things to see inside for example the Emerald Buddha image, the ordination hall and the gallery.	Present simple	Simple	<sup>4</sup> Pronoun+ V to be +Quantifier+Adj+N+ To +V+ Complement
Ah!There <b>are</b> 178 the mural painting[s] at the gallery.	Present simple	Simple	Similar to 1
There [They] <b>are located</b> opposite to the[/θɪ/] ordination hall.	Present simple	Simple	Similar to 2
They <b>were painted</b> in the reign of King Rama the first and <b>renovated</b> many times.	Past simple	Compound	Similar to 2



<p>Dusit Mahaprasat Throne Hall <b>is</b> very important because it [is] <b>used</b> for the annual consecration ceremony.</p> <p>First at 08 o'clock we <b>will depart</b> from The Grand Hotel Bangkok.</p> <p>Next, at 0930 we <b>will arrive</b> at Nakhon Pathom, <b>visit</b> The Golden Pagoda and <b>pay respect</b> to the scared Buddha image.</p>	<p>Present simple</p> <p>Future 'will'</p> <p>Future 'will'</p>	<p>Complex</p> <p>Simple</p> <p>Compound</p>	<p>Similar to 3</p> <p><sup>5</sup>Cohesive marker+Prep+Time+Future 'will'+Complement Similar to 5</p>
<p><b>Task type one (Low group):</b> Name it <b>is</b> the[di] Emerade [Emerald] Buddha Temple.</p> <p>Location it <b>is</b> <b>lotate[located]</b> in intent[inner] section of the Royal Land[Grand] Palade [Palace].</p> <p>Built, it <b>was build[built]</b> in the rain[reign] of King Rama the fird[first] in [pause] seventeen eighty two.</p> <p>First at egg [eight] o'clock we <b>will depart</b> from The Gand[Grand] Hotel Bangkok.</p> <p>Next at [pause], 09[pause] we will we <b>will [pause] alive[arrive]</b> at Nakhon Pathom [pause] <b>visit</b> [pause] Golden Pagose[Pagoda] and <b>pay repack[respect]</b> to the secard[sacred] Buddha imade [image].</p>	<p>Present simple</p> <p>Present simple</p> <p>Past simple</p> <p>Future 'will'</p> <p>Future 'will'</p>	<p>Simple</p> <p>Simple</p> <p>Simple</p> <p>Simple</p> <p>Compound</p>	<p>Similar to 1</p> <p>Similar to 2</p> <p>Similar to 2</p> <p>Similar to 5</p> <p>Similar to 5</p>
<p><b>Task type two (High group):</b> Please <b>don't take</b> photo in the ordition[ordination] hall because it <b>is</b> the the regulation.</p> <p>Please <b>don't make</b> noit[noise] when you [are] inside the ordition[ordination] hall because it <b>will bother</b> other people.</p>	<p>Present simple</p> <p>Present simple, Future 'will'</p>	<p>Complex</p> <p>Compound-complex</p>	<p><sup>6</sup>Adv 'Please'+ Auxiliary not+Conjunction+ Complement Adv 'Please'+ Auxiliary not +V+Adv clause+Conjunction+ Complement</p>

Please <b>be quite [quiet]</b> in the ordination hall because we <b>should [pay] respect</b> this place.	Present simple	Complex	Adv 'Please'+ V+Conjunction+ Complement
<p><b>Task type two (Low group):</b> You <b>should not put</b> shoes because [pause] it's impolite.</p> <p>You <b>should be</b> punctual because you <b>will miss</b> the bus</p> <p>Please <b>do not come late</b> because you will miss the bud [bus].</p>	Present simple,	Complex	Pronoun+Modal 'should' not +V+ Conjunction+ Complement
	Present simple, Future 'will'	Complex	Pronoun+Modal 'should'+V+ Conjunction+ Complement
	Present simple, Future 'will'	Complex	Similar to 6
<p><b>Task type three (High group):</b> Uhm Bangkok also <b>offer[s]</b> the best kind of food on the planet [pause] and when you <b>travel</b> in Bangkok you <b>must see</b> some activity in Bangkok uhm such as Khao San Road, Paragon uh movie at cinema theater, floating market.</p> <p><b>Don't worry</b> madam. I <b>will send</b> someone for giving some medicine for your son right now.</p> <p>And if your son <b>is</b> not better, <b>can</b> you <b>call</b> me back?</p> <p>OK uh I <b>will introduce</b> to uh tour program for today and [pause] we <b>will see</b> the sunset on the uhm behind the temple.</p>	Present simple	Compound complex	Noun+ Verb+ Complement+ Adv clause
	Present simple, Future 'will'	Simple	Auxiliary not+ Verb Sir/Madam.Pronoun+ Future 'will'+ Complement
	Present simple	Complex	Conjunction+If clause+Complement
	Future 'will'	Compound	Pronoun+ Future 'will'+ Complement
<p><b>Task type three (Low group):</b> I'm sorry sir and madam. I <b>will call</b> the driver right now and it <b>will never happen</b> again.</p> <p>OK. sir and madam. I <b>will shenk[change]</b> your new room for you.</p>	Present simple, Future 'will'	Compound	Pronoun+V to be+Adj+Sir/Madam+ Pronoun+ Future 'will'+ Complement
	Future 'will'	Simple	Exclamation 'OK'+ Sir/Madam. Pronoun+ Future 'will'+ Complement

I'm so sorry sir and madam because in the place that place it <u>close[s]</u> [to] inpruse[improve].	Present simple	Complex	Pronoun+V to be+Adj+Sir/Madam+ Conjunction+ Complement
--	----------------	---------	--

(The underlined parts represent verbs in tenses. Only a particular structure that frequently occurred in the scripts is presented in the table.)

In Task type one, present simple and past tense were found in the two proficiency groups while the future 'will' was mainly used in the sub-subsequent task. It is noted that the passive voice both in present and past tense was mainly used in this task type. Simple and complex sentences were also the primary constructions in this task type. For example, the student A1.1 used a simple construction "This is the Emerald Buddha Temple" to introduce the attraction, but used the complex sentence "Dusit Mahaprasat Throne Hall is very important because it [is] used for the annual consecration ceremony", to point out the importance of the attraction. This student also used compound construction "They were painted in the reign of King Rama the first and renovated many times" to explain the details of the wall gallery, and used the compound-complex sentence "The Emerald Buddha Temple was very important because Thai people believed that it was the most sacred place[s] in Thailand and the repository of spirits for all Thai people" to explain the importance of the site. The last type of sentence was mainly found in the high proficiency group at 43 frequencies whereas only 20 frequencies were used by the low proficiency group. The analysis showed a particular structure associated with Task type one as presented from pages 124 to 125.

In Task type two, present simple and future 'will' tenses were mainly used by the two proficiency groups. The analysis showed that complex was the prominent construction in this task type. It was noted that the compound-complex sentence was mainly found only in the high proficiency students' responses at 11 frequencies while only three frequencies were found in the low proficiency group's responses. The sample of complex construction was taken from the student B2.3 "Please do not come late because you will miss the bud [bus]" and the compound-complex construction was from the student A2.1 "Please don't make noit [noise] when you [are] inside the ordination [ordination] hall because it will bother other people". Examples of other

particular structures in this task type are presented in the excerpts from pages 125 to 126.

Similar to Task type two, present simple and future 'will' were used by the two proficiency groups of students in Task type three. Simple and compound constructions were mainly used in this task type by the two proficiency groups. Complex constructions were limitedly found in the two proficiency groups' responses at two frequencies for the low proficiency group whereas nine frequencies were found in the high proficiency group's responses. Compound-complex construction was only used by the high proficiency group (four frequencies). The low proficiency student B3.16 employed simple construction "OK sir and madam. I will shenk[change] your new room for you" to respond to the tourist complaint and the compound construction was from the student B3.10's excerpt as "I'm so sorry sir and madam. I will call the driver right now and it will never happen again". Compound-complex construction was very limitedly used by the high proficiency students and the example was from A3.5 as "Uhm Bangkok also offer[s] the best kind of food on the planet [pause] and when you travel in Bangkok you must see some activity in Bangkok uhm such as Khao San Road, Paragon uh movie at cinema theater, floating market". Examples of other particular structures in this task type are also presented from pages 126 to 127.

On the contrary, the salient grammatical difference in the responses of the two proficiency groups was in the range and complexity of the tenses and structures. The high proficiency group used present perfect (eight frequencies), past continuous (18 frequencies) and past perfect (four frequencies) tenses in Task type one. The high proficiency group used more types and more sophisticated sentences than the low proficiency group across the three task types, particularly on compound-complex sentences that were mainly used in the high proficiency groups' responses at 58 frequencies whereas 23 frequencies were used by the low proficiency group. Among the three task types, the difference was noted in the last task type in that four sentence types were found in the high proficiency group responses whereas three types were used by the low proficiency group. The examples of these grammatical differences between the two proficiency groups are presented from pages 124 to 127.

Furthermore, another salient grammatical difference was in accuracy with the low proficiency group making a majority of the errors, whereas these were limited in the high proficiency group's responses. These errors were the lack of verbs and wrong use of verb forms in the sentence structures. Details of the lack of verbs are presented below.

B1.6	Picture5: This is the Gand [Grand] Palace. Itch [It] wash [was] built by King Rama the 1 <sup>st</sup> in 1789. The Gand [Grand] Palace is very important because itch[it] ( <u>is</u> ) one of the location[s] inside the wall.
A2.8	Please you ( <u>be</u> ) quiet because [pause] you [pause] pay respect to the Buddha image.
A2.4	Please don't ( <u>take</u> ) a photograph the forbid and...
B2.9	Please ( <u>don't leave</u> ) belonging on the bus.
B2.12	Please ( <u>take</u> ) free service to JJ Mall near the tourist police office.
B3.9	I'm so sorry sir I will shenk[change] program tour for you right now. I will never ( <u>let it</u> ) happen again.

(The underline parts represent verbs in tenses.)

The error in the lack of verbs was mostly found in the low proficiency group's performances and it was prominent in Task type two. In task type one, the student B1.6 did not use 'is' to make the grammatical sentence. This error also occurred in the excerpt of A2.8 in that the infinitive 'be' was not used. Similarly, the student A 2.4 left the verb 'take' from the response; thus, making the sentence ungrammatical. For the low proficiency student B2.9, the modal verb with infinitive 'don't leave' was not used which caused inaccurate and incomprehensible sentences. The same error existed in the response of the student B2.12 in that 'take' was left out from the sentence. Similarly, the student B3.9 did not include 'let it' in the sentence and thus caused both an ungrammatical sentence and unclear meaning. The examples of wrong use of verb forms are given below.

B1.1	Picture4: This is the gallery. It was built in the reign of Kring [King] Rama [the] first. It wat [was] built by King Rama the first. It is <i>note</i> [noted] for its murals depicting the entire รามเกียรติ์. There are many interesting [things] to see for example the inside wall (was) <i>decorate</i> [decorated]...
B1.4	Picture4: There [They] are <i>locase</i> [located] oppo opposite the[di] ordination hall. There [They] were <i>panting</i> [painted] in the reign of King Rama the firt [first] and rensovate [renovated] many time[s].
A2.5	Please do not <i>climbing</i> [climb] Buddha image because it is impolite in Thai culture and dangerous.
A2.3	Please <i>sitting</i> [sit] politely in front of the Buddha sta ju [statue] because you should be respected [pay respect].
B2.4	You should not <i>climbing</i> [climb] because Thai people respect Buddha imade[images].
B2.7	You should <i>sitting</i> [sit] in front of the Buddha stature[statue] because Thai people respect the Buddha stature[statue].
B3.10	I'm so sorry sir and madam becaud [because] today that place <u>close (closes)</u> to improve.

(The italic words represent errors in grammar.)

The error in wrong use of verb forms was found across three task types and mostly found in the low proficiency group's responses. In Task type one, students used the wrong verb form, particularly past participle tense as in verbs 'noted' and 'decorated' in the excerpt of the student B1.1, and the verb 'located' and 'painted' in the student B1.4's excerpt. The error in Task type two was related to the incorrect use of present participle verb form after the modal verb which was noted in this task type. The students A2.5 and B2.4 incorrectly used verbs 'climbing' after the modal 'do not' and 'should not'. The student B2.7 also ungrammatically used 'sitting' after the modal 'should' whereas the student A2.3 used the similar verb form after 'please'. In Task type three, students misused present simple tense as in the verb 'closes' in the excerpt of the student B3.10. Additionally, some errors were noted in particular task types and the examples are given below.

Students	Excerpts	Types of grammatical errors				
		Adj.	Prep.	Pron.	Noun	Verb
		Wrong use adj to adv	Wrong use prep.	Wrong use of pronoun	No noun	Wrong use infinitive 'be'
A2.6	Please dress up <i>polite</i> [politely] because we should show pay respect this place.	✓				
B2.4	You should not sit <i>impolite</i> [impolitely] because [pause] it's the rule.	✓				
B1.9	<i>In</i> [At] [pause], 09[pause] thirty arrive at Nakhon Pathom [pause] visit [pause] the Golden Pagoda and pay respect to the secard[sacred] Buddha imade [image].		✓			
B2.16	<i>In</i> [At] ten thirty visit ทวารวดี Museum.		✓			
B2.19	Please <i>to</i> always carry the map because you will be get lost [pause] on the way. You should call <i>to</i> the tourist police when you get lost.		✓ ✓ ✓			
B1.4	Picture4: <i>There</i> [They] are locase [located] oppo opposite the[di] ordination hall. <i>There</i> [They] were panting [painted] in the of King Rama the firt [first] and rensovate [renovated] many time[s].			✓ ✓		
B3.8	I'm so sorry sir and madam. It will never happen again. I will check <i>it program</i> for you.			✓		

B3.17	I'm so sorry sir and madam because in the place <i>that place</i> it close inpruse [improve].			✓		
A1.7	Picture4: There are many <i>interesting</i> [thing] to see for example ...				✓	
B1.1	Picture1: ... There are many <i>interesting</i> [thing] to see for example the ordination [pause] hall and the Emerald Buddha imesh [image].				✓	
A2.5	Please <i>be keep</i> your wallet because it lost or stolen wallet.  Please <i>be call</i> police if you any problem stolen wallet lost variable[valuable] belongings because police can help you.  Please <i>be you have</i> hotel gard [card] because you can <i>going</i> anywhere					✓  ✓  ✓

(The italic word represents grammatical errors. Only the errors related to the topic are presented in the italic form.)

The first grammatical error was in the incorrect use of adjective and adverb and this error was noted in Task type two. The student A2.6 incorrectly used adjective 'polite' instead of 'politely' after the verb 'dress up'. Another example was from the student B2.4 that incorrectly used 'impolite' instead of 'impolitely' after 'sit'. The second error was in the wrong use of time preposition in Task type one and two. The student B1.9 incorrectly used preposition 'in' instead of time preposition 'at' whereas the student B2.16 added 'to' in the sentence. The student B2.19 also incorrectly added 'to' in the speech after the transitive verb 'call'. The third error was the wrong use of pronoun in Task type one and three. The student B1.4 incorrectly used indefinite pronoun 'there' instead of the plural pronoun 'they' for 'paintings'. The student B3.8 also incorrectly used both 'it' and 'program' in the sentence; and



B3.17 ungrammatically added both ‘that place’ and ‘it’ in the answer. The fourth type of error was a lack of a noun in the sentence in Task type one. The students A1.7 and B1.1 did not include nouns such as ‘things or places’ to make grammatical sentences. The last error was related to the wrong use of infinitive ‘be’ with other infinitives. This error was mainly found in the high proficiency students in Task type two. The student A2.5 ungrammatically added infinitive ‘be’ with the main verbs ‘keep’, ‘call’ and ‘have’ in the three sentences.

- **Language functions**

Language functions were investigated on the appropriate use of the language functions. In Task type one, the two proficiency groups similarly employed ideational function by explaining to the tourists about the attractions in Bangkok and the tour itinerary. For Task type two and three, manipulative function was found in giving suggestions to the tourist tasks. Similarly, in Task type three, manipulative function was mainly used to show regret and provide solutions to the tourists.

The reason for using manipulative function in the last task type may be that the two proficiency groups showed their regret to the tourists and directly offered solutions to each test task. Both showing regret and offering solutions were in manipulative function. The example of the use of the manipulative function from the high proficiency group is presented in the excerpt of the student A3.1. In contrast, the low proficiency group mainly used the expressions “I’m sorry sir and madam. It will never happen again” to respond to all test tasks as in the excerpt of the student B3.4, showing that the students may not understand the task and mainly relied on the memory to respond to the task. The examples of the use of language functions from the two proficiency groups are given in the following excerpts.

Students	Excerpts	Types of language functions
A1.3	<p>Picture1: We are now standing in front of the most sacred structure[s] in Thailand, the[/θɪ/] Emerald Buddha Temple. It was built in the reign of King Rama the first. It is the symbol of Thai nation and the repository of spirits for all Thai people.</p> <p>Task 2:</p> <p>Good morning ladies and gentlemen. I'm I'm Jariya. I would be your tour guide for the rest of your stay here. Now I'm [Ø]would like to tell you uh itinerary uh of Bangkok-Nakhon Pathom trip.First at 08 o'clock we will depart from the Grand Hotel Bangkok.Then, at 09:30 we will a arrive at Nakhon Pathom and we will visit The Golden Pagoda and pay respect to the scared Buddha image.After that, at 10:30 we will visited [visit] the ทวารวดี Museum.</p>	<p><b>Ideational:</b> explaining the construction at the Grand Palace and providing information about the tour itinerary</p>
B1.3	<p>Picture7: This is Dusit Mahaprasat Tone [Throne] Hall. Id [it] ead[is] conctructed[constructed] in 1786 by King Rama the firt [first]. It id [is] important becourd [because] it id [is] also use[used] for unnal [annual] consecvation [consecration] day ceremony. There are many interesthink [interesting] things to see inside such as the [pause] พระ[pause] ราชบัลลังก์ประดับมุก.</p> <p>Task 2:</p> <p>First, at egg [eight] o'clock we will depart from The Grand Hotel Bangkok.</p> <p>Next, at [pause] 09[pause]30 o'clock we will arrive at Nakhon Pathom [pause] visit [pause] The Golden Pagoda and pay respect to the sacred Buddha imade [image].Then at 1030 o'clock we will [pause] visit ทวารวดี Museum...</p>	

A2.9	<p>Please take off your shoes before entering in the ordination hall because it's more impolite than you take it.</p> <p>Please do not climb the Buddha image because we should pay respect [to] the Buddha image.</p> <p>Please beware [of] your pickpocket [pocket] when you [go] shopping.</p> <p>Please contact a tourist police if you [have] a problem.</p> <p>Please do not miss the bus because the bus leave[s] on time.</p> <p>Please take free transfer to BTS because it's fast and convenient.</p>	<p><b>Manipulative:</b> giving polite suggestion to the tourists about do's and don'ts at the ritual site and at the crowded attractions</p>
B2.20	<p><i>Do not</i> include the singlet and shorts because disrespectful.</p> <p><i>Do not</i> take the photograph because we should respect this place.</p> <p><i>Do not</i> use a loud voice [voice] because it's sacred [sacred] place.</p> <p><i>Do not</i> climbing the Buddha image [image] because you should respect the Buddha image [image].</p>	
A3.1	<p>Certainly, that's no problem. If you want we will go visit and shopping at the floating market.</p> <p>I'm sorry. I will give the medicine for your son right now. Your son shouldn't drink cool water. I will try my best.</p>	<p><b>Manipulative:</b> offering solutions to the tourists'</p>
B3.4	<p>I'm sorry sir and madam. I will change a room right now It will never happen again. I'm sorry sir and madam. It will never happen again.</p>	

(The italic words represent pragmatic mistake in language functions. Grammatical errors are not corrected.)

Concerning the difference in this LSP component, it could be found in the pragmatic mistake that was related to proficiency levels in the target language and the pragmatic competence on the degree of politeness. Similar to the previous LSP components, this mistake was mainly made by the low proficiency students as in the excerpt of B2.20. This student made the mistake by using the direct command 'Do not' with the tourists who were the audiences and the customers. In order to give the suggestions to the tourists, the pragmatic awareness on the degree of politeness is

required. Due to low proficiency levels in the target language and limited competence in the pragmatic knowledge, the low proficiency students may be less aware of the degree of politeness than the high proficiency students, and the low proficiency students may do not know that ‘Do not’ is considered the direct command, which should not be used with the tourists. For these reasons, their speech productions did not contain the appropriate degree of politeness, thus, resulted in the inappropriate use of the language functions, particularly in Task type two. As part of the construct by using the polite language, ‘Do not’ would be considered impolite and inappropriate.

- **Cohesion**

Cohesion was investigated on the types (connectors, relative pronouns and time sequence markers), and number of cohesive markers per response. It is measured by the accuracy and range of cohesive markers. From the three task types, the high proficiency groups employed similar types of cohesive markers (connectors, relative pronouns and time sequence markers) as their counterpart. Examples of the use of cohesive markers from the two proficiency groups in the three task types are presented as in the following excerpts.

Students	Excerpts	Types of cohesive markers	
		Connector	Time sequence
A1.1	<p>Picture1: This is the Emerald Buddha Temple. It was built in the reign of King Rama the first in 1782. The Emerald Buddha Temple was very important <b>because</b> Thai people believed <b>that</b> it was the most sacred place[s] in Thailand and the repository of spirits for all Thai people. There are many interesting things to see inside <b>for example</b> the Emerald Buddha image, the ordination hall and the gallery.</p> <p>Task 2: Bangkok-Nakhon Pathom <b>First</b> at 8 o'clock we will depart from The Grand Hotel Bangkok. <b>Next</b>, at 09:30 we will arrive at Nakhon Pathom and visit The Golden Pagoda <b>and</b> pay respect to the scared Buddha image. <b>Then</b>, at 10:30 we will visit พิพิธภัณฑ์ Museum. <b>Next</b>, at 11:30 we will [go] sightseeing</p>	Because, that, for example, and, such as	First, next, and, then,

	<p><b>and</b> buy a souvenirs at the Local market. <b>After that</b>, have lunch at ‘Ban Ruen Thai Restaurant’, a famous restaurant in Nakhon Pathom, <b>that</b> offers the delicious local dish[es] <b>such as</b> [pause]grilled[gill] river prawn[pround], spicy salad[pause] with fresh squid, fried chicken with Thai herbs, and spicy Nakhon Pathom soup at 12:45. <b>Then</b>, we will visit ‘Sanam[pause]chandra Palace’[place] at 13:45. And the <b>last</b> [lastly] we will depart from Nakhon Pathom at [pause] fifteen o’colck. <b>Finally</b>, we will arrive safely at The Grand Hotel Bangkok at [PAUSE] forty thirty fourteen thirty.</p>		<p>after that, that, lastly, finally</p>
B1.3	<p>Picture1: Did [This] id [is] the Imerald [Emerald] Buddha Temple. It is located [pause] in the eastern section of the royal Gand [Grand] Palate [Palace]. It wash [was] built in the rain [reign] of King Rama the firt [first] in 1782. Id [it] id [is] important <b>becausd</b> [because] it is believet [believed] to be the mode [most] sacresh [sacred] place in Thai [Thailand]. [pause] There are many interesthink [interesting] things to see inside <b>suck</b> [such] <b>ad</b> [as] the Emeralsh [Emerald] Buddha image, the ordination hall and the gallery.</p> <p>Task 2</p> <p><b>First</b>, at egg [eight] o’clock we will depart from The Grand Hotel Bangkok. <b>Next</b>, at [pause], 09[pause]30 o’clock we will arrive at Nakhon Pathom [pause] visit [pause] the Golden Pagoda and pay respect to the sacred Buddha imade [image].<b>Then</b>, at 10:30 o’clock we will [pause] visit ทวารวดี Museum. <b>After that</b>, at 11:30 o’clock [go] sightseeing and buy a souvenir at the Local market. <b>Then</b>, at 12:45 o’clock we will have lunch at ‘Ban Ruen Thai Restaurant’, a famous restaurant in Nakhon Pathom, <b>that</b> offers the delicious local dishes such as grill led [grilled] river prawn, spicy salad with fresh square [squid], fried chicken with Thai herbs, and spicy Nakhon Pathom soup. <b>Next</b>, [at] 13:45 o’clock we will visit ‘Sanam[pause]chandra Palade [Palace]’[pause] after that at [pause] fif fifteen o’colck depart from Nakhon Pathom. <b>Finally</b>, at [PAUSE] sixteen thirty o’clock arrive</p>	<p>Because, such as</p>	<p>First, next, then, that, after that, finally</p>

	safety safely at The Grand Hotel Bangkok.		
A2.7	You should take off your shoes <b>because</b> [pause] we should [pay]respect [to] this place <b>if</b> you wear shoes ,floors are dirty. You should not climb on the Buddha imade [image] because ritual site <b>and</b> Thai people uhm and Thai people pay respect [to the] Buddha imade [image].	Because, if , and , when	
A2.20	You should carry the map <b>when</b> you get loast [lost] it can help you.		
B2.5	You should call to the tourist police <b>because</b> you get the lot[lost] <b>and</b> have a problem.	Because, and , when	
B2.19	You should call to the tourist police <b>when</b> you get lost.		
A3.5	... Uhm Bangkok also offer the best kind of food on the planet [pause] and <b>when</b> you travel in Bangkok you must see some activity [activities]in Bangkok uhm such as Khao San Road, Paragon uh movie[s] at cinema theater, floating market.	When, if, and, that, because	
A3.16	Don't worry madam. I will send someone for giving some medicine for your son right now. And <b>if</b> your son is not better, can you call me back? I I will try my best to help your son.		
A3.7	OK uh I will introduce to uh tour program for today and [pause] we will see the sunset on the uhm behind the temple <b>and</b> for your information just now I will find your answer next time. OK sir I'm recommended the Dusit Princess Hotel <b>because</b> it has a beer garden and a folk song for you.		
B3.10	I'm so sorry madam I will shank[change] your order right now <b>and</b> it will never happen again. I'm so sorry sir and madam <b>becaud[because]</b> today that place close [is closed] to improve.	And, because	

(The bold font represents cohesive markers. Grammatical errors are not corrected)

The analysis from the frequency counts of cohesive markers across three task types and from the two proficiency groups' responses showed the salient feature in Task type one in which the time sequence cohesive devices were mainly used in this task type (314 frequencies) which may be from the task content and task requirement. The example was from the student A1.1's excerpt in that 'first, next, then, after that, lastly and finally' were used to explain about the sequence of the tour program. In Task type two, the connector 'because' (384 frequencies) was primarily used by the two proficiency groups to give the explanation about Thai etiquettes in the religious site and what the tourists should do at the crowded attractions as in excerpts of the students A2.7 and B2.5. In Task type three, 'and' and 'because' were mainly used by the two proficiency groups to respond to the tourists' enquiries and complaints in this task type (and=46 frequencies, because=15 frequencies) as presented in excerpts of students A3.5, A3.7 and B3.10. However, the connector 'if' was only found in the responses of the high proficiency group in this task type (six frequencies) as in the excerpt of the student A3.5.

On the contrary, the salient difference between the two proficiency groups was in the range of the cohesive devices in the responses used by the high proficiency group. High proficiency students employed twice as many cohesive markers as the low proficiency students. From the total of 1,425 cohesive markers, the high proficiency groups reported the total of 834 cohesive markers in their responses whereas 591 markers were found in the low proficiency group's performances. The salient difference was in Task type one and three. In Task type one, the high proficiency students used the total of 516 cohesive markers whereas 375 of them were found in the low proficiency students' responses. As for Task type three, the high proficiency group reported 77 cohesive devices while 20 of them were used by the low proficiency group. The two proficiency groups used almost similar frequencies of cohesive markers in Task type two. The high proficiency group used the total of 241 cohesive markers and 196 of them were used by the low proficiency group in this task type.

Additionally, the errors in limited or no use of cohesive markers that caused an unconnected speech were mainly found in the low proficiency group's responses. The examples are as follows.

B1.9: This is one day trip itinerary from Bangkok to visit Nakhon Pathom. The **first** in egg [eight] o'clock depart from The Grand Hotel Bangkok. In [pause], 09[pause] thirty arrive at Nakhon Pathom [pause] visit [pause] The Golden Pagoda and pay respect to the sacred[sacred] Buddha imade [image]. In ten thirty visit ทวารวดี [Tavaravati] Museum. In eleven thirty sightseeing and buy a souvenirs at the Local market...

B2.1: Going anywhere with hotel ไม่ไม่รู้จะอธิบายยังไงดี [Don't know how to explain].

Take the bag on the bud[bus].

OK to JJ Mall the tourist .

B3.18: OK will tow[tell] you now. Jack[Just] moment please. I[pause] call[pause] to[pause] the driver now.

Uh OK I will call the doctor now. Please [pause] give me the doctors.

OK I'm sorry. The place now for nice please.

From these excerpts, it was salient that all of the low proficiency students produced unconnected sentences. The student B1.9 used only one connector, the first, in the whole speech, whereas the students B2.1 and B3.18 did not show any evidence of using the cohesive devices.

- **Fluency**

Fluency was investigated on the appropriate use of tempo and pauses in the speech. The two proficiency groups similarly made the prominent errors in the inappropriate use of tempo and pauses that obstructed the flow of the responses, and caused difficulty in understanding the responses. This included repeating words or



phrases in the responses. However, all of the errors were mainly found in the low proficiency group's responses across the three task types. In each task type, the low proficiency students used very slow tempo in the responses with short pauses between words. They also used inappropriately long pauses in the middle and at the beginning of the responses that caused unconnected ideas, and directly affected the understanding of the responses. Another error was repeating words or phrases that obstructed the flow and understanding of the responses. The examples of the two proficiency groups are illustrated in the following excerpts. The bold font represents errors in pauses and repeating words and phrases; and the line represents slow tempo as the students were reading word by word. Examples of these errors are provided below.

Students	Excerpts	Types of errors		
		Inappropriate pauses	Repeating words or phrases	Reading word by word
A1.5	<p>Bangkok-Nakhon Pathom</p> <p>First, at eight o'clock <b>de_depart</b> from the Grand Hotel Bangkok.</p> <p>Then, at twelve forty-five o'clock have lunch at 'Ban Ruen Thai Restaurant', a famous restaurant in Nakhon Pathom, that offers the delicious local dish[es] such as grilled] river prawn[pround], spicy salad[<b>pause</b>] and fresh squid, fried chicken with Thai herbs, and spicy Nakhon Panom[Pathom] soup.</p> <p>Then, [<b>pause</b>] at [<b>pause</b>] <b>uh uh</b> สิบสาม at thirteen [<b>pause</b>] forty-five o'clock visit 'Sanamchandra Palace'. After that, at <b>fif_ fif_ fifteen o'clock</b> depart from Nakhon Pathom. Finally, at sixteen thirty arrive <b>safe_ safely</b> at the Grand Hotel Bangkok.</p>	<p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p>	<p>✓</p> <p>✓</p> <p>✓</p> <p>✓</p>	
B1.9	<p>Picture4: There_ are_one_ hundred_ seventy egg [eight]_ the mural painting[s] at_the gallery. There are_locase [located] <b>oppo_opposite</b> the ordination_ hall. There_ were_ panting</p>		<p>✓</p>	<p>✓</p>

	<p>[paintings] in_ the_ of King Rama the[<b>pause</b>] firt [first] and_ <b>resovate_renovate</b> [renovated] many_ time[s] [<b>pause</b>] the_stoley [story] of the panting [painting] was_about_is Ramayana or <b>Rama_</b> <b>Ramakien</b> the painting [s] are important_ becourd [because] [<b>pause</b>]they[<b>pause</b>]repace_[<b>reflect</b>] Ratanakosin_art.</p>	✓		
A2.12	<p>Please_ take_ off_ your_ shoes before entering because _to_ pay_ respect_ to_ the_ place_ you visiting [visit].</p>			✓
B2.6	<p>Please _you_ take_ off _your _shoes [<b>pause</b>] becaud [because] <b>becaud</b> [because]_not it is to _it_ is_ ruse[rule] มันเป็นกฎ.</p> <p><b>Please do not take</b> please_ do_ not_ take_ a_ photo_ becaud [because] it_ is_ the_ ruse[rule]</p> <p>Do_ not_ stop_ mouth _[<b>pause</b>] <b>becaud</b> [because] _<b>polite_</b> <b>becaud</b> [because]_ it_ is_ polite</p> <p>Please_ beware [<b>pause</b>]</p> <p>Tourist_police_ call_ becaud [because] [<b>pause</b>]</p>	✓	✓	✓
B3.13	<p>Just_ a moment_ please <b>I_ I</b> _change_ the_ room_ now.</p> <p>I'm_ sorry[<b>pause</b>] I_ will_ not_ let_ it_ happen again.</p>	✓	✓	✓
B3.18	<p>Uh OK_ will_ tow[tell]_ you_ now. Jack_[Just] moment_ please. I[<b>pause</b>] call[<b>pause</b>] to[<b>pause</b>] the driver now.</p> <p>OK_ I'm_ sorry [<b>pause</b>] Excuse_ me [<b>pause</b>] please.</p>	✓		✓

	Uh_ I'm_ sorry [pause] I [pause] will[pause] help the_ driver _now[pause] please.	✓ ✓		
--	---	--------	--	--

From the above excerpts, the majority of the students from the three task types produced the error in the reading words, particularly the low proficiency students which caused their speech to be too slow and difficult to understand. Another prominent error was in the inappropriate use of pauses which obstructed the flow of the speech and thus caused the speech to be hard to understand. The student B1.9 paused in the middle of the speech and did not stop or pause at the end of the sentence. “There\_ were\_ panting [paintings] in\_ the\_ of King Rama the[pause] firt [first] and\_ resovate\_rensovate [renovated] many\_ time[s] [pause] the\_stoley [story] of the panting [painting] was\_about\_is Ramayana or Rama Ramakien the painting [s] are important\_ becourd [because] [pause]they [pause] repace\_[reflect] Ratanakosin\_art”. There should be a pause at the beginning of the new sentence “the story of the painting” and “the paintings are important because” to separate the content in the speech and also to mark the information for the audience. For the error in repeating words or phrases, the majority of the students repetitively pronounced parts of the word or a whole word such as “fif fif fif fifteen o’colck” by the student A1.5, “oppo\_opposite” by the student B1.9, “Please do not take please\_ do\_ not\_ take” by the student B2.6 and “I\_ I \_change” by the student B3.13. Some students also repeated phrases as in the excerpt of the student B2.6, “Please do not take please\_ do\_ not\_ take”.

- **Content knowledge**

Content knowledge was investigated on types of content knowledge found in the speech and measured by the accuracy and completion of the information in responding to the test tasks. As part of the specific feature of an LSP test in the form of test content, the analysis showed that the two proficiency groups similarly reported a specific type of content knowledge associated with a particular task type. In Task type one, the content knowledge related to Thai architectural structures, particularly temples and palaces, Thai arts, Thai history of the temples and palaces and Buddhism

was found. In Task type two, Thai cultural knowledge on do's and don'ts at the religious sites was noted. This task type also included do's and don'ts at the tourist attraction. In Task type three, content knowledge was mostly related to the problem-solving in tourism-oriented situations, particularly in dealing with tourists' enquiries, and responding to tourists' complaints. The following excerpts show different types of content knowledge used by the two proficiency groups and grammatical errors are not corrected.

Students	Excerpts	Types of content knowledge
A1.3	<p>Picture3: This is the highlight of our trip today. It is called the ordination hall. It was built in the reign of King Rama the 1<sup>st</sup> [pause]. It was built by King Rama the 1<sup>st</sup> and renovated by King Rama the third.</p> <p>What is the interesting is that there are uh there is the Emerald Buddha image enshrined in it. There are many other interesting things such as the garudas holding nagas at the base of the hall. The mural paintings and the two large standing crown Buddha images.</p> <p>Picture 6: We are now standing in front of the พระที่นั่งจักรีมหาปราสาท group. This building was constructed by King Rama the fer_fifth to commemorate the centenary of the Chakri dynasty. The eastern porch has the reception room where portraits of kings of the Chakri dynasty from King Rama the first to King Rama the seventh [pause] are displayed. In the west porch is the hall where portraits of queens of the King Rama the fourth, King Rama fifth and King Rama the seventh are displayed. In the rare [rear] center is the Chakri Throne Room. Here the king received ambassadors on the occasion of the presentation of the credentials.</p>	<p>History about the temple</p> <p>Thai visual arts and statues</p> <p>History and details about the palace</p>
A2.5	<p>Please do not take your shoes in the ordination hall because it is impolite.</p> <p>Please do not climbing Buddha image because it is impolite in Thai culture and dangerous.</p> <p>Please do not sit the impolite posture in front of Buddha satstus [statue] because it is bad in Thai</p>	<p>Thai cultural knowledge about the etiquette at the religious site</p>

	<p>culture.</p> <p>Please be keep your wallet because it [can be] lost or stolen wallet.</p> <p>Please be punctual because maybe you miss [the] trip. Please be [use] free service to [at] JJ Mall near the tourist police office because it save money.</p>	<p>Tourism knowledge at the tourist attractions</p>
<b>A3.5</b>	<p>Don't worry madam. I will send someone for giving some medicine for your son right now. And if your son is not better, can you call me back? I will try my best to help your son.</p> <p>I'm sorry to hear that madam uh I don't know the smoke come[s] to your room and uh I will contact housekeeping [housekeeper] for your room right now. And if it's not better you can call me back later. I will not let it happen again.</p>	<p>Knowledge on dealing with tourists' enquiries</p> <p>Knowledge on responding to tourists' complaints</p>

In contrast, the difference was in the errors in the incorrect and incomplete information of the content knowledge that were mostly found in the low proficiency groups' responses across the three task types. Some of these errors were limitedly found in the high proficiency students' performances. The following excerpts illustrate the error in incorrect content knowledge from the two proficiency groups in the three task types.

Test takers	Excerpts	Types of incorrect content knowledge			
		Number	Etiquette	Do's and don'ts at the attractions	Attraction recommend
A1.2	<p>Picture 2: This is the Umerald [Emerald] Buddha image. It was made from jade. It is in meditation posture. <i>It is [pause] 48 [pause] point 3 centimet [centimeters] wide and high.</i> It has [pause] 2 seasonal costumes made in the reign of King Rama the 1<sup>st</sup>. It is now one for enchane [enshrining] inside the ordination hall.</p>	✓			

B1.8	Picture5: This is the Gand [Grand] Pace [Palace]. Itch [it] wash [was] built in the rain [reign] of King Rama the firt [first] in 2782 the Gand [Grand] Pace [Palace] id [is] important becourd [because] it id [is] resident [residence] of King Rama there are many intersthink [interesting] things to see inside for example the พระมหามณเฑียร group, the พระมหาราชวัง group, the พระที่นั่งจักรีมหาปราสาท group the Monpiman Monpiman [Boromphiman] Mansion group and คิวาลัย garden group.	✓			
A2.12	Please do not take a photograph because <i>you are not allowed to take it in this private place.</i>		✓		
B2.4	You should not take a photo because <i>it's peach copyright [it is a copyright place].</i>		✓		
B2.20	You should not take a photo on because <i>it's danger.</i>		✓		
B2.7	You should left your belonging because <i>it saves your life.</i>			✓	
B2.11	<i>You should not take your back[bag] on the bus because to lost [it can get lost]on the way.</i>			✓	
A2.6	Please beware the pickpocket because <i>we will steal your pocket.</i>			✓	
A2.14	Please be careful for your wallet because <i>it's rules.</i>			✓	
B2.2	You should take free the transfer to BTS because <i>it's the government policy.</i>			✓	
B2.14	Please take the free transfer to [the] car because <i>it is the</i>			✓	

	<i>promotion.</i>				
B2.16	Please <i>your sit</i> [take your seat] <i>on the bus because it is promotion[pause] to you</i> จบหรือยังไม่. นะ.			✓	
B2.11	You should be on time because <i>work late.</i>			✓	
B2.14	Please do not miss the bus because <i>it is dangerous.</i>			✓	
A3.7	OK sir I'm recommended the <i>Dusit Princess Hotel</i> because it has a bear garden and a folk song for you.				✓
B3.3	Sure I am agree with you going to <i>Phimai History Park. It's so beautiful. Phimai History Park [and]its building with Chaivoraman the seventh.</i>				✓
B3.20	OK sir and madam. <i>I will shenk[change] a new program</i> for you.				✓

(The italic font represents errors in content knowledge and grammatical errors are not corrected.)

All of the students from the two proficiency groups in Task type one made the same mistake about the number. The mistake on the first excerpt was about the size of the Emerald Buddha Image and the correct information was 48.3 centimeters wide and 66 centimeters high. The second excerpt showed the mistake about the year of construction of the Grand Palace which was 1782 not 2782.

In Task type two, the error in wrong information of the content knowledge was about the Thai cultural etiquette inside the temple. The majority of the students gave the wrong reason for prohibiting visitors from taking the photos inside the ordination hall by saying that it was because of the copyright, private place and danger. The correct reason should be “You should not take a photo because it is considered to be disrespectful manner in Buddhism beliefs”. Another salient error was about the do’s and don’ts at the tourist attractions related to giving suggestions for

taking care of the belongings and wallets, taking free transfer and taking the bus on time.

For Task type three, the prominent error was about recommending the famous attractions in Bangkok for the tourists. The students suggested the wrong places by giving the places from their hometown or providing wrong information as in the excerpt of the student B3.3.

Additionally, the errors in insufficient content knowledge and repetitive use of the similar information to respond to the different prompts were mainly found in Task type two and three, especially in the low proficiency students' responses. The examples of the insufficient content knowledge error are illustrated as follows.

<b>Students</b>	<b>Insufficient content knowledge</b>	<b>Type of content knowledge</b>
A2.6	Please be quiet because this place want[s] the peace. You should not clam Buddha imade [image] because ritual [it is sacred].	Etiquette at the temple
A2.11	You should check two things before get [getting] off the bus.	Do's and don'ts at the tourist attractions
B2.1	Going anywhere with hotel ไม่ไม่ผู้จะอธิบายยังงี้ Take the bag on the bud [bus]. OK to JJ Mall the tourist.	Do's and don'ts at the tourist attractions
B2.6	Please sit down Please sit down. Please beware [pause]. Tourist police call becaud [because][pause]. Please cat[carry] map becaud [because] not get lot[ lost]. Please do not [leave your] belonging becaud [because].	Etiquette at the temple  Do's and don'ts at the tourist attractions
B2.8	Please do not beware [pause].	Do's and don'ts at the tourist attractions
B2.9	Please take off your shoes.	Etiquette at the temple
B2.12	Please [take] free service to JJ Mall near the tourist police office.	Do's and don'ts at the tourist attractions
B2.13	Please sitting [sit] in front of the [pause]Buddha status[statue]. I must find tourist police becaud [because].	Etiquette at the temple Do's and don'ts at the tourist attractions



B2.14	Please do not climbing [climb] because it is [sacred image].	Etiquette at the temple
A3.14	I will check it now. I will not lets it [let it happen again].	Dealing with wrong room request
A3.18	I will help you now. Just [a] moment a please.	Dealing with sick tourists
B3.6	I'm so sorry and madam I will never [let it] happen again.	Dealing with incomplete tour program
B3.7	I'm so sorry sir and madam. It will never [let it] happen again.	Dealing with late bus
B3.11	I'm sorry very much [very sorry]. I will do it batter[better]	Dealing with incomplete tour program
B3.14	I'm so sorry sir and madam. It will never happen again.	Dealing with incomplete tour program

From the excerpts, most of the errors were from Task type two, particularly in the etiquette at the religious site and do's and don'ts at the attraction. The majority of the students from the second task type did not provide adequate information for the suggestions. The example was the excerpt of B.26 on sitting politely in front of the Buddha statue which was the way to show respect in Thai culture. For the do's and don'ts at the tourist attraction, the students also mentioned only the suggestions without giving sufficient information. The examples were from the excerpts of A2.11 and B2.1 that were related to suggesting the tourists about taking their valuable belongings with them, while two students did not give any reasons for their suggestions. It would be clearer for the tourists with the additional information such as for the safety of their valuable things.

Similar to Task type two, the students in Task type three did not provide adequate information for the tourists and hardly showed any content knowledge. This task type dealt mostly about the problem-solving skill of the tour guide; therefore, adequate information was required to deal with all the situations. The examples were from the excerpts of A3.14 and A3.18 in which the students should add more information to resolve the problem such as "I will contact the front office" for the first excerpt and "I will take your son to the hospital" for the second situation. For B3.6, B3.11 and B3.14, the situation dealt with tourists' complaints for the incomplete tour

program and additional information would have made the situation better such as “I will check again and if we have enough time, I will try to take all of you to other places”. Additionally, the examples of error in repetition are given below in italics.

A2.13: You should not wear spaghetti shirt [singlet] because *it's the rule*.

You should not take a photo in the temple because *it's the rule*.

You should not speak louder [loudly] inside the ordination hall because *it's the rule*.

You should not climbing [climb] on the Buddha image because *it's the rule*.

B2.1: Please do not wear spaghetti shirt [singlet] and shorts in the the tepen [temple] because *it is noly [holy]*.

Please do not [pause] speak [pause] power [loudly] because *it's [the] rule*.

Please do not [pause] walk in [climb] the [pause] in the Buddha because *it's [the] rule*.

Please do not [pause] sitting [sit] in the front of your the Buddha because because *it is noly [holy]* ใ้ยั คั้ันนั้จะผิิด ว่าคอะไ้รนะเนี้ย.

A3.6: *I'm sorry I will check it now. I will not let it happen again* [wrong room request].

*I'm sorry I will check it now. I will not let it happen again* [sick tourist].

*I'm sorry I will check it now. I will not let it happen again* [lost the wallet].

*I'm sorry I will check it now. I will not let it happen again* [incomplete tour program].

B3.11: *I'm sorry I will check it now* [wrong room request].

*I'm sorry I will check it now. I will not let it happen again* [sick tourist].

All of the students repetitively used the same information to respond to the different prompts which showed both inaccurate and irrelevant content knowledge. In Task type two, the students employed the sentence “it is the rule” and “it is holy” to respond to a number of prompts and some were inaccurate information. The first example was from A2.13’s excerpt on the situation related to the prohibition on climbing the Buddha image which was not the rule, but it would be disrespectful in Thai belief. Another example was from B2.1’s excerpt on the situation of keeping quiet inside the ordination hall. There was no rule in Thai culture to prohibit people to speak inside this religious place, but it would either disturb other people or would not show respect to the sacred ritual site.

In Task type three, the students repetitively used “I’m sorry I will check it now” and “I’m sorry I will help you now” to respond to almost all the prompts. In some situations, these sentences were considered irrelevant content knowledge; consequently, it might cause the speech to be hard to understand. The examples were from the excerpts of A3.6 and B3.11 with the situation related to a sick tourist. The students employed “I’m sorry I will check it now. I will not let it happen again” which was irrelevant because they did not have to say sorry to the tourist and they could not guarantee that the unexpected situation like this would not happen.

To summarize, the analysis revealed some similarities and differences of all the LSP components from the two proficiency groups’ performances across the three task types. It was noted that proficiency levels strongly affected the difference in the LSP productions between the two proficiency groups in terms of accuracy, range, complexity and appropriateness in all components. Most of the errors in each component were mainly found in the low proficiency students’ responses while a wider range in vocabulary, tenses, grammatical structures and cohesive markers were used in the high proficiency students’ responses. The high proficiency students also employed more compound-complex sentences than their counterpart. They used more

types and more appropriate language functions in their performances than the low proficiency students. The high proficiency group produced more fluent speeches and had more accurate pronunciation and content knowledge than the low proficiency group.

In contrast, similarities in each LSP individual component of the two proficiency groups' performances were found, and some were noted in a particular task type. In pronunciation, the similarities were in the incorrect pronunciation of words, ending of words and wrong use of stress across the three task types. In addition, the two proficiency groups made some errors that were prominent in a particular task type: no intonation, incorrect pronunciation of Thai words and incorrect pronunciation of consonant clusters in Task type one and two. For vocabulary, the two proficiency groups similarly employed tourism-related technical terms in Task type one. They made the typical error in the inaccurate use of vocabulary in the three task types while the error in the use of generic terms was noted in Task type two. Moving to grammar, present simple and future 'will' tenses were used across the three task types whereas past simple tense was mainly used in Task type one. Simple and compound constructions were used in Task type one and three while complex was mainly used in Task type one and two. The analysis indicated particular structures with each task type. The two proficiency groups made typical errors in the lack of verbs and wrong use of verb forms in the sentences. They made errors that were prominent in particular task types such as in the misuse of adjectives in Task type two, preposition in Task type one and two, pronoun in Task type one and three, no noun in Task type one and infinitive 'be' in Task type two. Additionally, the analysis showed that the two proficiency groups used particular language functions in a certain task type. They employed the ideational function in Task type one while manipulative function was used in Task type two and three. One pragmatic mistake in the language functions was found in Task type two in the use of the direct command with the tourists. For cohesion, the two proficiency groups used similar types of cohesive markers (connectors, relative pronouns and time sequence markers) in their responses. The two proficiency groups used connectors and relative pronouns across the three task types while time sequence markers were mainly found

in their performances in Task type one. In terms of the use of the range of cohesive devices, the high proficiency group used twice as many cohesive markers as the low proficiency group in their responses. The two proficiency groups made typical errors in limited or no use of cohesive markers. For fluency, the two proficiency groups made typical errors in inappropriate pauses, repeating words or phrases, and reading word by word across the three task types. The results also showed a particular content knowledge in relation to the task type. The two proficiency groups similarly employed history and art content knowledge in Task type one, Thai etiquettes knowledge and tourism knowledge in Task type two, and knowledge on dealing with tourists' enquiries, and responding to tourists' complaints in Task type three. They made typical errors in inaccurate information, while some errors were noted in a particular task type: insufficient content knowledge and use of repetitive information in Task type two and three.

#### **4.2. Research question 2: What are the students' attitudes towards the WBST-EFT?**

Research hypothesis 4: There is no significant difference in the attitudes towards the WBST-EFT in the two proficiency groups at the .05 level.

##### **4.2.1. Students' attitudes on the four aspects of the WBST-EFT**

To answer the second research question, the mean scores and standard deviations were obtained from the attitudes towards the WBST-EFT online questionnaire from the high and low proficiency students. The scores were interpreted into degrees of attitudes: 4=*Strongly agree*, 3=*Agree*, 2=*Disagree*, and 1=*Strongly disagree*. Additionally, some items of the questionnaire were presented in the negative statements, 1.7, 1.8, 3 and 4.2, and the reverse score calculation was, therefore, used with these items. The results are presented in Tables 4.24 to 4.28.

**Table 4.24: Mean differences of the students' attitudes on the four aspects of the WBST-EFT**

Attitudes towards the Web-based Speaking Test in English for Tourism (WBST-EFT)	Proficiency levels				df	t	p
	High		Low				
	$\bar{x}$	SD	$\bar{x}$	SD			
1. Overall usefulness	3.31	0.37	3.21	0.39	118	1.44	0.19
2. Appropriateness of time for preparation and response formulation	3.27	0.58	3.25	0.60	118	0.15	0.87
3. Task difficulty	3.65	0.73	3.67	0.60	118	0.14	0.89
4. Interface design	3.37	0.40	3.26	0.49	118	1.32	0.19
<b>Total</b>	<b>3.40</b>	<b>0.35</b>	<b>3.35</b>	<b>0.41</b>	<b>118</b>	<b>0.76</b>	<b>0.45</b>

The findings from Table 4.24 reveal that there is no significant difference at .05 level between the two proficiency groups in the four aspects,  $t(118) = .76$ ,  $p > .05$ ; and their total means are not much different ( $\bar{x}_H = 3.40$ ,  $SD = .35$ ,  $\bar{x}_L = 3.35$ ,  $SD = .41$ ), so the fourth hypothesis is accepted. In other words, the two proficiency groups similarly agreed with the statements, and they had positive attitudes from the high mean scores towards the WBST-EFT in four aspects. In addition, the high proficiency group has slightly higher mean scores than the low proficiency group in most of the aspects with the exception to the task difficulty. The reverse score was applied with this aspect. The low proficiency group outnumbered their counterpart ( $\bar{x}_H = 3.65$ ,  $SD = .73$ ,  $\bar{x}_L = 3.67$ ,  $SD = .60$ ), indicating that the two proficiency groups similarly agreed that the test was not too difficult for them. The students' views on the individual aspect of the test are presented in the following part.

#### **4.2.2. Students' attitudes on the individual aspect of the WBST-EFT**

The results of the individual items are presented in Tables 4.25 to 4.28, and the following table illustrates the results of students' views towards the overall usefulness of the test.

**Table 4.25: Mean differences of the students' attitudes on the usefulness of the WBST-EFT**

Overall usefulness	Proficiency levels				df	t	p
	High		Low				
	$\bar{x}$	SD	$\bar{x}$	SD			
1.1 I feel the WBST-EFT provided me the adequate opportunity to demonstrate both of my strengths and weaknesses on speaking ability.	3.21	0.55	3.16	0.61	118	0.46	0.64
1.2 I think the raters listening to my response via the WBST-EFT will get an accurate idea of my speaking ability in tourism context as stated in the course syllabus.	3.41	0.59	3.21	0.69	118	1.70	0.09
1.3 I felt at ease taking the WBST-EFT.	3.20	0.57	3.08	0.64	118	1.04	0.29
1.4 The instructions are clear and easy to follow.	3.38	0.69	3.13	0.79	118	1.84	0.06
1.5 The introduction part is useful because it gives me the example and chance to practice the test.	3.43	0.59	3.23	0.67	118	1.72	0.87
1.6 The tasks and situations on the WBST-EFT are appropriate and simulate the real world tasks.	3.31	0.67	3.20	0.68	118	0.94	0.34
1.7 The test taking procedures are too sophisticated for me and require proficiency in computer.	3.23	0.89	3.15	0.80	118	0.54	0.59
1.8 The web-based test is not an appropriate test for speaking ability.	3.27	0.63	3.53	0.83	118	-1.97	0.051
<b>Total</b>	<b>3.31</b>	<b>0.37</b>	<b>3.21</b>	<b>0.39</b>	<b>118</b>	<b>1.44</b>	<b>0.19</b>

Table 4.25 shows no significant difference at .05 level between the two proficiency groups' views on all of the sub-items of the overall usefulness of the test,  $t(118) = 1.44$ ,  $p > .05$ ,  $\bar{x}_H = 3.31$ ,  $SD = .37$  and  $\bar{x}_L = 3.21$ ,  $SD = .39$ , indicating that they

similarly agreed with all statements in this aspect. Almost all of the higher means are from the high proficiency group except the last sub-item which is from the low proficiency group ( $\bar{x}_H=3.27$ ,  $SD=.55$  and  $\bar{x}_L=3.53$ ,  $SD=.61$ ). The reverse score was used with this item and it can be interpreted that both groups agreed that the web-based test was appropriate for speaking assessment. Moreover, the reverse score was also used with the item 1.7 and the scores can be interpreted that both proficiency groups agreed that the test taking procedures in the WBST-EFT were not too sophisticated for them. The following table displays the results of the appropriateness of time for preparation and response formulation.

**Table 4.26: Mean differences of the students' attitudes on the appropriateness of time for preparation and response formulation of the WBST-EFT**

Appropriateness of time for preparation and response formulation	Proficiency levels				df	t	p
	High		Low				
	$\bar{x}$	SD	$\bar{x}$	SD			
2.1 The preparation time for the test is adequate.	3.30	0.67	3.28	0.69	118	0.13	0.89
2.2 The time allowed for response formulation for the test is appropriate.	3.23	0.70	3.22	0.67	118	0.13	0.89
<b>Total</b>	<b>3.27</b>	<b>0.58</b>	<b>3.25</b>	<b>0.60</b>	<b>118</b>	<b>0.15</b>	<b>0.87</b>

From Table 4.26, there is no significant difference between the two proficiency groups in this aspect,  $t(118) = .15$ ,  $p > .05$ ,  $\bar{x}_H=3.27$ ,  $SD=.58$  and  $\bar{x}_L=3.25$ ,  $SD=.60$ , indicating that the two proficiency groups similarly agreed with all statements in this aspect. All the p values are significant at .05 level. As for the sub-item, the two proficiency groups' means are not much different. The high proficiency group has slightly higher mean scores in the appropriateness of time for preparation ( $\bar{x}_H=3.30$ ,  $SD=.67$ ) than the low proficiency group ( $\bar{x}_L=3.28$ ,  $SD=.69$ ). Similarly, the higher mean score in response formulation ( $\bar{x}_H=3.23$ ,  $SD=.70$ ) is from the high proficiency group whereas the lower one is from the low proficiency group ( $\bar{x}_L=3.22$ ,  $SD=.67$ ). In other words, both proficiency groups similarly agreed that the time for



the preparation and response formulation for the test was appropriate. The students' views on the task difficulty are presented as follows.

**Table 4.27: Mean differences of the students' attitudes on the task difficulty of the WBST-EFT**

Task difficulty	Proficiency levels				df	t	p
	High		Low				
	$\bar{x}$	SD	$\bar{x}$	SD			
3. I think the tasks in the WBST-EFT are too difficult.	3.65	0.73	3.67	0.60	118	-0.14	0.89
<b>Total</b>	<b>3.65</b>	<b>0.73</b>	<b>3.67</b>	<b>0.60</b>	<b>118</b>	<b>0.14</b>	<b>0.89</b>

Table 4.27 shows that there is no significant difference at .05 level between the two proficiency groups in this aspect,  $t(118) = .14, p > .05$ , showing that they had similar attitudes on the difficulty of the test tasks. The reverse score was used with this item; and this means that the two proficiency groups similarly agreed that the test tasks were not too difficult ( $\bar{x}_H = 3.65, SD = .73$  and  $\bar{x}_L = 3.67, SD = .60$ ). The following table presents the students' views on the interface design of the test.

**Table 4.28: Mean differences of the students' attitudes on the interface design of the WBST-EFT**

Interface design	Proficiency levels				df	t	p
	High		Low				
	$\bar{x}$	SD	$\bar{x}$	SD			
4.1 I think the navigation button, icon (e.g. recording and move buttons), tool bar and controls are easy to use.	3.23	0.62	3.28	0.70	118	-.42	0.68
4.2 The terminology used in the test is hard to understand.	3.48	0.85	3.38	0.78	118	.67	.51
4.3 Each page layout contains appropriate detail, clear title and is easy to read.	3.32	0.68	3.17	0.70	118	1.2	0.23
4.4 I think the text, font size and color used in the test are appropriately designed.	3.30	0.65	3.25	0.65	118	.70	0.49

4.5 The multimedia (e.g. video clips & audio files) used in this test are appropriate.	3.42	0.53	3.23	0.70	118	1.62	0.11
4.6 The multimedia help me understand the prompt better and do not take too long to download.	3.38	0.64	3.27	0.63	118	1.03	0.32
4.7 I could use the help facilities (e.g. pop up explanation) while I was taking the test.	3.35	0.73	3.25	0.63	118	0.80	0.42
<b>Total</b>	<b>3.37</b>	<b>0.40</b>	<b>3.26</b>	<b>0.49</b>	<b>118</b>	<b>1.32</b>	<b>0.19</b>

Table 4.28 demonstrates no significant difference at .05 level between the two proficiency groups' views on the interface design of the test,  $t(118) = 1.32, p > .05$ ; and they had almost the same total means ( $\bar{x}_H = 3.37, SD = .40$  and  $\bar{x}_L = 3.26, SD = .49$ ). In other words, both proficiency groups similarly agreed with all statements of the interface design of this online speaking test. The high proficiency group has slightly higher mean scores than the low proficiency group in almost all of the aspects with the exception to the item 4.1 where the low proficiency group has slightly higher mean scores than the high proficiency group ( $\bar{x}_H = 3.23, SD = .62$  and  $\bar{x}_L = 3.28, SD = .70$ ), showing that the two proficiency groups similarly agreed that the icon and navigation button and tool bars were easy to use. The following part displays the content analysis of the open-ended part of this online questionnaire.

#### 4.2.3. Content analysis of the open-ended part

Referring to the first open-ended question on the students' attitudes about the strengths and weaknesses of the WBST-EFT, the two proficiency groups thought that the strengths of this LSP online speaking test were on the use of multimedia in the test task presentation, situations and tasks on the test and the relaxation on test taking procedures. 'High' and 'Low' represent the proficiency level of the students and the number in the parentheses indicates the assigned task type. The information in the brackets was added by the researcher for the clarity of the description.

As for the use of multimedia in the test task presentation, motion pictures and video clips were mentioned by the two proficiency groups. The high proficiency group thought that the motion pictures helped them to recall the information on the

attractions, historical background and important details on the test tasks. A high (1) student wrote that “I used pictures to remember the information of Wat Prakaew [The Emerald Buddha Temple] and Pra Tee Nang Chackri Mahaprasat [Chackri Mahaprasat Throne Hall], especially the history and background of the place[s]”. This was also referred by one low (1) student saying that “Picture[s] is [are] good for memory. I memory [remembered] many thing[s]”. High (2) and low (2) students used the pictures to memorize all the concepts and Thai cultural prohibitions. A high (2) student wrote that “Pictures help me remember the concept about do’s and don’ts in Thailand”. Low (2) student referred that “I remember[ed] the information about please do and don’t from the pictures. It helps a lot”. In addition, the video clips simulated real world situations and tasks were specified by high (3) and low (3) students in the way that they got the clear pictures of the situation. They thought of themselves as the tour guides who dealt with the tourists in different scenarios. High (3) student referred that “Clips made me look like the real guide”, “I can understand the situation more from the clips” and “I think that I am the real tour guide in the clip and help the tourists”. The low (3) student agreed that “Clips are good for understanding the situations” and “I like clip[s] because [I] look like the tour guide”.

Moving to situations and tasks, both groups thought that the situations and tasks were relevant to the actual tasks that would be performed by the professional tour guides. One of the high proficiency groups wrote that “I look like the real guide from the test task. The situations are good example[s] and I can use my knowledge in reality”. Some low proficiency students wrote that “I like [the] situation [because] I [can] help tourist[s]”, “I use knowledge [to] explain the tourist about Thai culture [as if] I am [were] the guide”. The high proficiency groups also mentioned that the situations were not too difficult as one of them said “It doesn’t [isn’t] difficult to understand the situations”.

The relaxation on the test taking procedures was also pointed out by the two proficiency groups as the strength of the WBST-EFT. Both proficiency groups mentioned that they could take the test without stress because there were no teachers who observed them during the test taking process. In the traditional face-to-face speaking test, a human interlocutor, who is normally the course lecturer, is required.

However, the interlocutor is not needed in this technology-integrated test. One of the high proficiency groups wrote that “It isn’t serious because you don’t have to speak in front of the teacher”. This is similar to the low proficiency groups saying that “No [without the] teacher [the test is] not [so] serious”. The high proficiency groups liked the preparation time of the test because they could gather the information to respond to the prompt. High (1) and (2) students referred that “I have time to prepare the test” and “I can think of the information to answer the test”.

As for the weaknesses of the WBST-EFT, the two proficiency groups mentioned about the technical problems, particularly on the recording system and the information requirement from the prompts. The high proficiency groups said that “The record is difficult to use”, “Recording system is quite hard”, “[It is] hard to save the voice”, “Sometimes, it's uncomfortable [not convenient] to entrance [enter] and save it” and “Sometimes, there was no voice in the record”. The low proficiency groups agreed with the high proficiency groups and they said that “[I] don’t like [to] save [the] file”. The two proficiency groups mentioned on the second weakness that “Too many information to answer”, “Answering information is too long” and “Speak too long [I have to describe a lot] with lots of information”.

Concerning the second question on what the students liked and disliked most about the WBST-EFT, the two proficiency groups liked the use of multimedia in the test task presentation and the test taking procedures without teacher. The two proficiency groups said that “I like movement picture[s] with sounds”, “I like the presentation for [the] test with video clips. It’s real” and “I like [the] speaking test. No teacher in front of you”. However, they disliked the sound recording and information requirement of the test. They mentioned that “I don’t like the system of recording sound. It[was] not complete”, “I think that it's uncomfortable [not easy] to save it” and “The test ask[ed] too many things”.

In summary, the two proficiency groups generally had similar views towards the WBST-EFT regarding the overall usefulness, appropriateness of time for preparation and response formulation, task difficulty and the interface design. Their views were found to be positive on the four aspects of the WBST-EFT from the high

total means of the four aspects. The content analysis also showed that the two proficiency groups liked the multimedia and relaxation in the test taking procedures which were regarded as the strengths of this online LSP test. The correspondence between situations of the test tasks and the real world tasks was considered an additional strength of this instrument. However, the technical problem, especially on the sound recording system was the weakness of this technology-based test.

#### **4.3 Research question 3: Are there any differences in types and frequency of speaking test taking strategies used by high and low proficiency students in doing the WBST-EFT?**

To answer the third research question, frequency and percentage of the reported strategies between the two proficiency groups were compared. Percentage relative to the total number of reported strategies was calculated to obtain the most frequently used strategies among the total reported strategies. Percentage relative to each type of strategies was also used to compare the most frequently reported sub-strategies within the individual category of the strategy. The results are presented into two parts; the first part is the comparison between the proficiency levels and the reported strategies, and the second part is the comparison among the task types, proficiency levels and strategies.

### 4.3.1 Comparison of proficiency levels and reported strategies

#### 4.3.1.1 Comparison of proficiency levels and total reported strategies

**Table 4.29: Proficiency levels and total reported strategies**

Proficiency levels		Types of strategies			Total
		Cognitive	Metacognitive	Communication	
<b>High</b>	Frequencies	87	50	23	<b>160</b>
	Percentage	33	19	9	<b>61</b>
<b>Low</b>	Frequencies	55	27	21	<b>103</b>
	Percentage	21	10	8	<b>39</b>
<b>Total</b>	<b>Frequencies</b>	<b>142</b>	<b>77</b>	<b>44</b>	<b>263</b>
	<b>Percentage</b>	<b>54</b>	<b>29</b>	<b>17</b>	<b>100</b>

Table 4.29 displays that the high proficiency group employed almost twice as many strategies as the low group at 61% comparing to 39% from the low proficiency group. Among the three types, Cognitive is the most frequently used strategy by the two proficiency groups at 54%, followed by Metacognitive at 29% and Communication at 17%. The obvious difference is in Cognitive strategy that was frequently used by the high proficiency group at 33% and by the low proficiency group at 21%. However, the high and low proficiency groups frequently employed almost similar percentage of Communication strategy at 9 and 8. Tables 4.30 to 4.32 display the individual sub-category reported by the two proficiency groups.

### 4.3.1.2 Comparison of proficiency levels and individual sub-category

**Table 4.30: Sub-categories of Cognitive strategy**

Profi. levels		Cognitive			
		Selecting	Comprehend	Storing memory	Retrieval
<b>High</b>	% relative to total number of strategies reported	6.46	8.74	7.98	9.88
	% relative to each type of strategy	11.97	16.19	14.78	18.30
<b>Low</b>	% relative to total number of strategies reported	2.28	8.74	7.22	2.66
	% relative to each type of strategy	4.22	16.19	13.38	4.92
<b>Total</b>	<b>% relative to total number of strategies reported</b>	<b>8.74</b>	<b>17.48</b>	<b>15.20</b>	<b>12.54</b>
	<b>% relative to each type of strategy</b>	<b>16.19</b>	<b>32.38</b>	<b>28.16</b>	<b>23.22</b>

Table 4.30 shows that the two proficiency groups frequently employed Comprehending (32.38%), Storing memory (28.16%), Retrieval (23.22%) and Selecting (16.19%) respectively. The high proficiency group used almost four times as many Retrieval strategies as the low group (18.30% and 4.92%). Similarly, Selecting strategies were used by the high proficiency group almost three times as many strategies as the low proficiency group (11.97% and 4.22%). However, the two proficiency groups equally employed Comprehending strategy at 16.19%. In addition, they used Storing memory at almost the same percentage at 14.78 for the high proficiency group and at 13.38 for the low proficiency group. The following table illustrates the sub-categories of Metacognitive strategy used by the two proficiency groups.

**Table 4.31: Sub-categories of Metacognitive strategy**

Proficiency levels		Metacognitive		
		Goal setting	Assessment	Planning
<b>High</b>	% relative to total number of strategies reported	5.70	6.08	7.22
	% relative within the category	19.48	20.77	24.67
<b>Low</b>	% relative to total number of strategies reported	1.90	4.56	3.80
	% relative within the category	6.49	15.58	12.98
<b>Total</b>	<b>% relative to total number of strategies reported</b>	<b>7.60</b>	<b>10.64</b>	<b>11.02</b>
	<b>% relative within the category</b>	<b>25.97</b>	<b>36.35</b>	<b>37.65</b>

Table 4.31 illustrates that the two proficiency groups frequently used Planning (37.65%), Assessment (36.35%) and Goal setting (25.97%) orderly. Among the three strategies, Goal setting is the strategy that differs the most between the two proficiency groups. The high proficiency group employed almost three times as many Goal setting strategies as the low proficiency group (19.48% and 6.49%). Planning was used by the high proficiency group twice as many strategies as the low proficiency group (24.67% and 12.98%). The least difference in this category is Assessment which was used by the high proficiency group at 20.77% and by the low proficiency group at 15.58%. Sub-categories of Communication strategy used by the two proficiency groups are displayed in the following table.



**Table 4.32: Sub-categories of Communication strategy**

<b>Profi. levels</b>		<b>Communication</b>						
		<b>Achievement (68%)</b>					<b>Avoidance (32%)</b>	
		<b>Appro.</b>	<b>Para.</b>	<b>Word coin.</b>	<b>Restruc.</b>	<b>Code switc.</b>	<b>Topic avoid.</b>	<b>Con. aband</b>
<b>High</b>	% relative to total number of strategies reported	-	3.42	0.76	1.52	0.38	1.52	1.14
	% relative within the category	-	20.45	4.54	9.09	2.27	9.09	6.81
<b>Low</b>	% relative to total number of strategies reported	1.90	3.42	-	-	-	1.14	1.52
	% relative within the category	11.36	20.45	-	-	-	6.81	9.09
<b>Total</b>	<b>% relative to total number of strategies reported</b>	<b>1.90</b>	<b>6.84</b>	<b>0.76</b>	<b>1.52</b>	<b>0.38</b>	<b>2.66</b>	<b>2.66</b>
	<b>% relative within the category</b>	<b>11.36</b>	<b>40.90</b>	<b>4.54</b>	<b>9.09</b>	<b>2.27</b>	<b>15.90</b>	<b>15.90</b>

Table 4.32 shows that the two proficiency groups employed more than twice as many Achievement strategies as Avoidance (68% and 32%). Among the seven sub-categories, Paraphrase is the most frequently used strategy by the two proficiency groups at 40.90%. It is followed by Topic avoidance (15.90%), Conversation abandoning (15.90%), Approximation (11.36%), Restructuring (9.09%), Word coinage (4.54%) and Code switching (2.27%) respectively. Approximation was only used by the low proficiency group which is the most differently reported sub-

categories between the two proficiency groups. Word coinage, Restructuring and Code switching were solely employed by the high proficiency group to solve the target language difficulty and to continue their responses, while the low proficiency group would abandon the conversation. Paraphrase was equally used by the two proficiency groups at 20.45%. Topic avoidance was more frequently employed by the high proficiency group (9.09%) than the low proficiency group (6.81%). In contrast, Conversation abandoning in which 9.09% was used by the low proficiency group and 6.81% was employed by the high proficiency group. The comparison of the task types, proficiency levels and types of strategies is illustrated in the following table.

#### 4.3.2 Comparison of task types, proficiency levels and reported strategies

##### 4.3.2.1 Comparison of task types, proficiency levels and total reported strategies

**Table 4.33: Comparison of the task types, proficiency levels and total reported strategies**

Task types	Proficiency levels	Types of strategies			Total	
		Cognitive	Metacognitive	Commun.		
1	High	Frequencies	32	16	5	53
		Percentage	12.16	6.08	1.90	20.14
	Low	Frequencies	14	6	8	28
		Percentage	5.32	2.28	3.04	10.64
	Total frequencies		46	22	13	81
	Total percentage		17.48	8.36	4.94	30.78

<b>2</b>	<b>High</b>	Frequencies	28	16	8	52
		Percentage	10.64	6.08	3.04	19.76
	<b>Low</b>	Frequencies	18	4	7	29
		Percentage	6.84	1.52	2.66	11.02
	Total frequencies		46	20	15	81
	Total percentage		17.48	7.60	5.70	30.78
<b>3</b>	<b>High</b>	Frequencies	27	18	10	55
		Percentage	10.26	6.84	3.80	20.90
	<b>Low</b>	Frequencies	23	17	6	46
		Percentage	8.74	6.46	2.28	17.48
	Total frequencies		50	35	16	101
	Total percentage		19	13.30	6.08	38.38
<b>Total</b>	<b>Total frequencies</b>	<b>142</b>	<b>77</b>	<b>44</b>	<b>263</b>	
	<b>Total percentage</b>	<b>54</b>	<b>29</b>	<b>17</b>	<b>100</b>	

Table 4.33 shows that the two proficiency groups frequently used most of the three strategies in Task type three at 38.38%; and they employed the same amount of strategies for both Task type one and two at 30.78%. However, the percentage of the reported strategies in each Task type is not much different.

Considering the use of the individual strategy in relation to each task type, Cognitive is the most frequently reported strategy in the first three ranks at 19% in Task type three and at 17.48% in Task type one and two. It is followed by Metacognitive that was frequently used in Task type three at 13.30%, in Task type one at 8.36% and in Task type two at 7.60%. For the last strategy, Communication

was reported in Task type three at 6.08%, in Task type two at 5.70% and Task type one at 4.94% orderly.

Concerning the difference in the use of each strategy in the three task types by the high and low proficiency groups, Cognitive was differently employed in Task type one (12.16% and 5.32%). Another difference is in Metacognitive strategy in Task type two which was used by the two proficiency groups at 6.08% and 1.52% respectively whereas Communication strategy was differently reported in Task type one (1.90% and 3.04%).

Regarding Task type one, the difference is in Cognitive. The high proficiency group reported more than twice as many Cognitive strategies as the low proficiency group (12.16% and 5.32%). The only strategy that was used more by the low proficiency group is Communication. It was used at 3.04% by the low proficiency group and at 1.90% by the high proficiency group. The high proficiency group reported the use of Metacognitive strategy three times more than that of the low proficiency group (6.08% and 2.28%).

As for Task type two, the most differently reported strategy in this task is Metacognitive strategy. The high proficiency group reported four times as many Metacognitive strategies as the low proficiency group (6.08% and 1.52%). It is followed by Cognitive strategies that were used by the high proficiency group almost twice as many strategies as the low proficiency group (10.64% and 6.84%). Communication strategies were reported almost at the similar amount by the high proficiency group at 3.04% and by the low proficiency group at 2.66%.

In Task type three, the high proficiency group reported almost as similar amount of Metacognitive strategies as the low proficiency group (6.84% and 6.46%). The high proficiency group reported almost twice as many Communication strategies as the low proficiency group (3.80% and 2.28%). Cognitive strategies were employed by the high proficiency group at 10.26% and by the low proficiency group at 8.74%; thus, it is the most differently reported strategy in this task. The table below points out the sub-categories of Cognitive strategy employed by the six groups.

### 4.3.2.2 Comparison of task types, proficiency levels and individual sub-category

**Table 4.34: Sub-categories of Cognitive strategy used by high and low proficiency groups**

Task types	Proficiency levels	Cognitive				
		Selecting	Compre.	Storing memory	Retrieval	
1	High	% relative to total number of strategies reported	1.90	3.42	3.42	3.42
		% relative within the category	3.52	6.33	6.33	6.33
	Low	% relative to total number of strategies reported	0.76	1.52	1.90	1.14
		% relative within the category	1.40	2.81	3.52	2.11
2	High	% relative to total number of strategies reported	2.28	2.28	2.28	3.8
		% relative within the category	4.22	4.22	4.22	7.04
	Low	% relative to total number of strategies reported	-	3.04	3.04	0.76
		% relative within the category	-	5.63	5.63	1.40

<b>3</b>	<b>High</b>	% relative to total number of strategies reported	2.28	3.04	2.28	2.66
		% relative within the category	4.22	5.63	4.22	4.92
	<b>Low</b>	% relative to total number of strategies reported	1.52	4.18	2.28	0.76
		% relative within the category	2.81	7.74	4.22	1.40
<b>Total</b>	<b>% relative to total number of strategies reported</b>	<b>8.74</b>	<b>17.48</b>	<b>15.2</b>	<b>12.54</b>	
	<b>% relative within the category</b>	<b>16.17</b>	<b>32.36</b>	<b>28.14</b>	<b>23.20</b>	

Regarding the four sub-categories, Comprehending is the most frequently reported strategy in the three task types by the two proficiency groups (32.36%). It is followed by Storing memory (28.14%), Retrieval (23.20%) and Selecting (16.17%).

In Task type one, the high proficiency group reported three times as many Retrieval strategies as the low proficiency group (6.33% and 2.11%) which is the strategy used most differently in this task type. On the contrary, the least difference is in Selecting which was used by the high proficiency group at 3.52% and by the low proficiency group at 1.40%. Storing memory was employed by the high proficiency group almost twice as the low proficiency group (6.33% and 3.52%).

Similar to Task type one, Retrieval strategy is the most differently reported strategy in Task type two. The high proficiency group employed five times as many Retrieval strategies as the low proficiency group (7.04% and 1.40%). In contrast, the low proficiency group employed more Comprehending and Storing memory strategies

than the high proficiency group (5.63% and 4.22%). However, Selecting was not reported by the low proficiency group in this task type.

As for Task type three, the two proficiency groups frequently reported similar amount of Storing memory at 4.22% while the most difference is in Retrieval. The high proficiency group reported more than three times as many Retrieval strategies as the low proficiency group (4.92% and 1.40%). The high proficiency group also employed almost twice as many Selecting strategies as the low proficiency group (4.22% and 2.81%). However, the low proficiency group outnumbered the high proficiency group in the use of Comprehending strategies at 7.74% to 5.63%. The following table displays the sub-categories of Metacognitive strategy reported by the high and low proficiency groups.

**Table 4.35: Sub-categories of Metacognitive strategy used by high and low proficiency groups**

Task types	Proficiency levels		Metacognitive		
			Goal setting	Assessment	Planning
1	High	% relative to total number of strategies reported	1.9	1.52	2.66
		% relative within the category	6.49	5.19	9.09
	Low	% relative to total number of strategies reported	-	1.52	0.76
		% relative within the category	-	5.19	2.59

<b>2</b>	<b>High</b>	% relative to total number of strategies reported	1.52	2.28	2.28
		% relative within the category	5.19	7.79	7.79
	<b>Low</b>	% relative to total number of strategies reported	-	0.76	0.76
		% relative within the category	-	2.59	2.59
<b>3</b>	<b>High</b>	% relative to total number of strategies reported	2.28	2.28	2.28
		% relative within the category	7.79	7.79	7.79
	<b>Low</b>	% relative to total number of strategies reported	1.9	2.28	2.28
		% relative within the category	6.46	7.79	7.79
<b>Total</b>	<b>% relative to total number of strategies reported</b>		<b>7.6</b>	<b>10.64</b>	<b>11.02</b>
	<b>% relative within the category</b>		<b>25.96</b>	<b>36.34</b>	<b>37.64</b>

Table 4.35 shows that Planning is the most frequently reported strategy by the two proficiency groups (37.64%), followed by Assessment (36.34%) and Goal setting (25.96%) in all the three task types.

In Task type one, the two proficiency groups reported the similar amount of Assessment strategies at 5.19%, but the low proficiency group did not report any Goal setting strategies at all. Therefore, it is considered as the most differently used strategy among the three task types. In addition, the high proficiency group employed Planning strategies three times more than the low proficiency group (9.09% for the



high proficiency group and 2.59% for the low proficiency group). This is the most difference in the reported strategy among the three task types.

For Task type two, the low proficiency group did not report any Goal setting strategies whereas they were used by the high proficiency group at 5.19%. It is considered the most differently reported strategy in this task. The high proficiency group employed Assessment and Planning strategies three times more than the low proficiency group (7.79% and 2.59%).

In the last task type, the two proficiency groups employed the same amount of Planning and Assessment strategies at 7.79%. The high proficiency group slightly used higher amount of Goal setting strategies than the low proficiency group (7.79% and 6.46%). Table 4.36 shows the sub-categories of Communication strategy used by high and low proficiency groups.

**Table 4.36: Sub-categories of Communication strategy used by high and low proficiency groups**

Task types	Profi. levels	Communication							
		Achievement					Avoidance		
		Appro.	Para.	Word coin.	Rest.	Code switc.	Topic avoid	Con. aband	
1	High	% relative to total number of strategies reported	-	0.38	-	-	0.38	-	1.14
		% relative within the category	-	2.27	-	-	2.27	-	6.81
	Low	% relative to total number of strategies reported	0.76	1.14	-	-	-	0.76	0.38
		% relative within the category	4.54	6.81	-	-	-	4.54	2.27

<b>2</b>	<b>High</b>	% relative to total number of strategies reported	-	1.52	-	-	-	1.52	-
		% relative within the category	-	9.09	-	-	-	9.09	-
	<b>Low</b>	% relative to total number of strategies reported	0.38	1.52	-	-	-	0.38	0.38
		% relative within the category	2.27	9.09	-	-	-	2.27	2.27
<b>3</b>	<b>High</b>	% relative to total number of strategies reported	-	1.52	0.76	1.52	-	-	-
		% relative within the category	-	9.09	4.54	9.09	-	-	-
	<b>Low</b>	% relative to total number of strategies reported	0.76	0.76	-	-	-	-	0.76
		% relative within the category	4.54	4.54	-	-	-	-	4.54
	<b>Total</b>	% relative to total number of strategies reported	<b>1.90</b>	<b>6.84</b>	<b>0.76</b>	<b>1.52</b>	<b>0.38</b>	<b>2.66</b>	<b>2.66</b>
		% relative within the category	<b>11.35</b>	<b>40.90</b>	<b>4.54</b>	<b>9.09</b>	<b>2.27</b>	<b>15.90</b>	<b>15.90</b>

Table 4.36 demonstrates that Paraphrase (40.90%) is the most reported sub-category of Communication strategy in the three task types. It is followed by Topic avoidance (15.90%) and Conversation abandoning (15.90%), Approximation (11.35%), Restructuring (9.09%), Word coinage (4.54%) and Code switching (2.27%) respectively.

In Task type one, the most difference in reported strategy is in Approximation and Topic avoidance which were solely employed by the low proficiency group at 4.54%. In contrast, Code switching was frequently and solely reported by the high proficiency group at 2.27%. The low proficiency group reported three times as many Paraphrase strategies as the high proficiency group (6.81% and 2.27%) while it is the reverse in the use of Conversation abandoning strategies ( 2.27% and 6.81%).

For Task type two, Topic avoidance is the most differently reported strategy in this task type. The high proficiency group frequently used four times as many Topic avoidance strategies as the low proficiency group (9.09% and 2.27%), but the two proficiency groups reported similar amount of Paraphrase strategies at 9.09%. Both Approximation and Conversation abandoning were solely reported by the low proficiency group at 2.27%. None of the two proficiency groups reported on Word coinage, Restructuring and Code switching.

In Task type three, Restructuring is the most differently reported strategy in this task and it was solely reported by the high proficiency group at 9.09%. Both Approximation and Conversation abandoning were only reported by the low proficiency group at 4.54% whereas the reverse is in Word coinage and it was only used by the high proficiency group at 4.54%. The high proficiency group also reported twice as many Paraphrase strategies as the low proficiency group (9.09% and 4.54%). None of the two proficiency groups reported using Code switching and Topic avoidance strategies. The following part presents the results of the content analysis from the verbal reports.

### 4.3.3 Content analysis from the verbal reports

This section shows the content analysis from the verbal reports on the strategies used by the 18 students in attempting the three task types of the WBST-EFT. The students' identification is coded by letter 'H', representing 'the High proficiency group', and letter 'L', representing 'the Low proficiency group'. The first number indicates the task type and the second number refers to the number of the reporter. The brackets represent the information added by the researcher for explicit understanding of the speech and '\*' represents the strategy that is not in the taxonomy. The categorization of the strategies in this study is based on Swain et al. (2009: 68-69) (See the full coding scheme in Appendix G).

#### 1. Communication strategy

Communication strategy deals with conscious planning to solve linguistic difficulty during communication. There are two strategies in this category: Achievement and Avoidance strategies.

1.1. Achievement strategies are used when the students face the communicative problem due to the lack of language knowledge. Six sub-categories are included in this category: Overgeneralization/morphological creativity, approximation, paraphrase, word coinage, restructuring and code switching. However, none of the students reported using overgeneralization or morphological creativity in doing this LSP online speaking test. The following reports present details of the five sub-categories.

1.1.1. Approximation strategy is used when the students cannot think of specific words which are mostly the technical terms. They employ more generic terms for the unknown words. This strategy was only used by the low proficiency group in the three task types. Some examples are given below.

Researcher: So, can you tell me what you were thinking when you were doing this task?

L1.1: [I] talked in general like I used “Place” instead of “พระที่นั่ง [Pavilion]” because I could not remember that vocabulary. I could not think of it.

L2.1: [I] used the vocabulary like “Clothes” for “กางเกงขาสั้น [shorts].” I could not remember that word.

1.1.2. Paraphrase strategy is used in situations when the students replace near synonym words to the unknown words. Similar to approximation, paraphrase is mostly employed in need of the technical terms and it was reported by all six groups of the students across the three task types (presenting tourism-related information, giving polite suggestions to the tourists and responding to tourists’ enquiries and complaints). The students replaced the difficult technical terms and employed the words that had the closest meaning that they could think of or the words that they learned from the class. For example,

H1.1: I substituted the difficult vocabulary and used the word that I knew and had the closest meaning.

H2.3: [I] used the vocabulary that I learned in the class for the unknown words in the picture. For example, “A place to see the Buddha image” was used for “พระอุโบสถ [the ordination hall]”.

H3.3: [I] tried to find the near synonym word like “Bus to BTS” for “รถรับส่งผู้โดยสาร [Transfer bus]”.

L1.3: [I] could not think of any words. [I] used the one that [I] could think of. [I] used “Sitting Buddha image” for “พระพุทธรูปปางมารวิชัย [Subduing Mara Buddha image]”.

L2.2: Sometimes [I] used the near synonym word. For example [I] used “Bad sitting position” to substitute “นั่งไม่สุภาพ [Sitting in an impolite position]”.

L3.2: [I] used easy and known vocabulary that [I] could remember at that time. The example was [when the tourist wanted to] change the room [I] used “Hotel person at the counter” for “พนักงานจองโรงแรม [Reservationist]”.

1.1.3. Word coinage strategy is reported when the students create the new words for the unknown words. It was only reported by the high proficiency group in doing Task type three (responding to tourists’ enquiries and complaints tasks). The example is given below.

H3.1: [I] created the new word. [I] did not know the word “คนขับรถทัวร์ [Bus driver]” because there was no preparation time. [I] needed to answer right the[a] way so [I] used “Chauffeur driver”.

1.1.4. Restructuring strategy is employed in the case when the students use the different structures to convey the similar meaning of the message and most of the time they substitute less complex structures to the more complex ones. This strategy was only used in Task type three by the high proficiency group. For instance,

H3.2: [I] used easy sentences that [I] could think of to answer [the question].

H3.3: [I] used the closest meaning sentence like “Sorry sir, I hope it’s not [doesn’t] happen again” for “I’m sorry sir I won’t let it happen again”.

1.1.5. Code switching strategy was only found in Task type one (presenting tourism-related information tasks) by the high proficiency group. The students used the native language words or phrases to replace the target language when they lacked of the target language knowledge. For example,

H1.1: [I] answered in Thai because I could not come up with any English words. [I] used “the พระที่นั่ง [the Throne Hall]” to answer.

1.2. Avoidance strategies are used when the students lack control over the target language. The strategies are made up of three sub-categories: topic avoidance, conversation abandoning and semantic avoidance. There was no report on using the last sub-category. Details of sub-categories are presented below.

1.2.1. Topic avoidance strategy involves avoiding the unknown topics. This strategy was reported by the low proficiency group in Task type one and two. It was mainly reported by the high proficiency group in Task type two, but none of this strategy was used in Task type three. For example,

L1.2: If [I] cannot remember the information [I] skip it for example, who was the constructors?

H2.2: [I] skipped the topic and continued with the new picture.

L2.2: [I] skipped the details on [the reason] why it was prohibited to wear [a] sleepless shirt inside the temple.

1.2.2. Conversation abandoning strategy is used when the students move to the next prompt by leaving the message incomplete. It was mostly used in Task type one by the two proficiency groups, but was only used by the low proficiency group in Task type two and three. Following are some examples.

H1.3: [I] gave short answer and skipped to the next item. It was better than did not say anything at all. “This is Chrackri Mahaprasart and...”.

L1.3: [I] skipped to the next question but I did say something. “พระอุโบสถ [The ordination hall] is...”.

L2.2: [I] skipped to the next one if [I] could not answer; for example, “Please don’t take off your shoes because....”.

L3.2: [I] skipped. For example, in the case that the tour program was changed. [I] could not come up with any reasons. I could not think of anything.

## **2. Cognitive strategy**

Cognitive strategy deals with manipulating the target language to understand and produce the language. Four sub-categories are included as follows.

2.1. Selecting or attending strategy is used to direct the attention to a specific feature of the task. The students focused on the different and prominent parts

of the test tasks. For example, majority of the students focused on the salient features of the pictures in the three task types. They also focused on the important information and key words in Task type one and two. For Task type three, the students emphasized on instructions, types of questions from tourists' speech, situations, vocabulary and tourists' gestures. Some examples are given below.

*Focusing on the salient features of the pictures*

H1.1: [I] looked at the picture, time and places and set my own answer. [I] arranged the information [which one should] come before or after to answer the question based on the existing information.

L1.1: [I] looked at the picture [and figured out] where the place was.

H2.3: [I] mainly looked at the pictures and thought of the vocabulary from the class. For example, the picture of do not wear [not wearing] shoes [I] recognized the salient feature of each picture. The picture with prohibition usually has a cross.

L3.1: [I] focused on ...., pictures and captured on tourists' speech...

*Focusing on important information and key words*

H1.3: [I] focused on the important part and answered in my own words. [To explain] วัดพระแก้ว [the Emerald Buddha Temple], [I] used "This is the Emerald Buddha Temple".

L1.1: ... [I] looked for the key words. [I] tried to find the focus of the place [so that I could] speak a lot. For another task [I] emphasized on time and place".

H2.2: [I] focused on the importance of the task .... For example, if [I] want to warn the tourist for not doing something [I] used "You should not climbing [climb] the Buddha image".



*Focusing on the instructions*

H3.1: [I] read the instructions of the task and tried to capture on what the tourists wanted.

*Focusing on the types of questions from tourists' speech*

H3.2: [I] listened and tried to figure out the type of questions whether it was requesting for help or giving [a] suggestion.

*Focusing on situations*

H3.3: [I] focused on the places [and] captured on what the tourists in the [video] clip wanted. For example, [if it was] the request we must help them.

L3.1: [I] focused on the situation... and captured on tourists' speech. I listened to find out where the situation took place....

*Focusing on vocabulary*

L3.3: [I] listened carefully to the vocabulary. [Tourists] spoke too fast [and I] could not catch up.

*Focusing on the tourists' gestures*

L3.1: ... including observed [observing] all the tourists' gestures.

2.2. Comprehending strategy deals with seeking ways to understand the task, and there are seven sub-categories: clarifying information, L1 translating, inference, analyzing, reasoning, \*imagination and \*summarizing. The students employed different types and number of sub-categories, particularly analyzing the strategy that was reported by all six groups across the three task types. However, L1 translation strategy was mainly used in Task type two and three by the two proficiency groups. Similarly, imagination was limitedly used by the high proficiency group in Task type one and three. Examples are given below.

*Clarifying information*

H1.3: [I] tried to understand the instructions from the task. For example, with the set time [I had to] arrange the answer, [I] emphasized on listing word orders such as “Next, Then [and] After that”.

*L1 translating*

H2.3: ...[I] translated [the instructions] into Thai for some pictures.

H3.1:...[I] translated [the conversation] into Thai to better understand [the situation].

L3.3: [I] translated into Thai ...

*Inferences*

H2.3: [I] used the information from the prompt to understand the task and figured out whether it was “Should do” or “Should not do”...

*Analyzing*

H1.1: ... [I] analyzed the task on what [I] needed to do. [Figure out whether] it was easy or difficult in order to prepare [for] the information to answer the task.

L1.3: [I] analyzed the instructions and prepared for the information. For example, [I] thought of how to use “First, Next and Finally” with the information.

H2.1: ... analyzing the task [to find out what I] should do, should not do or just the warning.

L2.3: [I]... analyzed on what the task asked; whether it was “Should” or “Should not do” such as “You should not take a photo because it is a rule”.

H3.1: ...[I] followed up the conversation in the task; and analyze whether it was complaint or sorry.

L3.3:...[I] analyzed the situation on what the tourists wanted...

*Reasoning*

L1.1: [I] looked for the importance of the place and think of the information and the reason.

H2.1: Then, [I] thought about the reason to answer by looking at the picture like in the tourist police, [I said] “If you have problem, please call this number. We can help you”.

H3.1...Then, [I] think of the reason to resolve the situation.

L3.3...and [I] thought of the right answer with the supporting reasons.

*Imagination*

H1.1: First, we needed to think that we are the real tour guide to understand the task more...

H3.1: We must think that we are the actual tour guide to better understand the situation...

*Summarizing*

H1.2:...[I] summarize [the information] and figure out the picture and the details in order to explain [the picture] uhm [I] really tried to solve the problem in the situation.

2.3. Storing memory strategy deals with finding ways to memorize information to respond to the test tasks. It involves repeating what is read or heard, associating the existing information, linking new information with previous knowledge, summarizing the L2 information, using imagery to generate, understand or remember the information, \*memorizing linguistic features, \*focusing on main idea of the information,\*categorizing information and \*memorizing situations. Using imagery and repeating what is read or heard are the most frequently reported sub-categories by almost all of six groups of the students whereas none of them reported

using any linking new information with the previous knowledge. For some strategies namely: focusing on main ideas, categorizing information and summarizing L2 information were reported only in Task type one by the two proficiency groups. Similarly, memorizing situation strategy was reported only by the high proficiency group in attempting Task type three. Following are some examples.

*Repeating on what is read or heard*

H1.2: [I] .... reread it [the information] such as which reign was [the place] constructed? [So that] I could answer the question...

L1.2: [I] tried to reread and remember the information from the pictures and vocabulary.

L2.1: [I] reread the instruction until I could remember everything ....

L3.2: ...and [I] reread the instruction.

*Associating the existing information*

H3.1: [I] tried to use what I learned in the class to answer [the question] by relating to the information from the tasks.

L2.1: ... and [I] tried to use the information that I have learned to answer [the task]. I link it with the detail in the task.

L3.1: [I] used the knowledge from the class to respond.

*Summarizing the L2 information*

L1.3: [I] briefly summarized the English information in order to be easily understood.

*Using imagery to generate, understand or remember the information*

H1.1: ...Or [I] memorized the picture to recall the vocabulary and put them [it] into sentences in my own word.

H2.3: Most of the time [I] used pictures to memorize. They were very beneficial in memorizing the information and [I] memorized the outstanding features of the picture.

L2.3: [I] memorized the picture to give the answer in English. For example, [I said] “You should take off your shoes because it is a rules” or “You should sit polite [politely because] it is a rules”. [I] mostly used very easy vocabulary that [I] could remember.

L3.2: [I] tried to use the picture to memorize ....

*Memorizing linguistic features*

H1.1: Most of the time, [I] memorized the vocabulary from the class and made sentences from that...

H3.2: ...[I] memorized the [sentence] structure.

*Focusing on main idea of the information*

H1.2: [I] focused on the gist, ... The example was “The Emerald Buddha Image is [uh][uhm] the symbol of Thai nation[uh] many people go to worship [uh]”. Something like this.

*Categorizing information*

L1.1: [I] categorized the piece of information to easily memorize the information such as what was the name? Where was it? Why was it important?

*Memorizing situations*

H3.3: [I] used the situation to help me memorizing the information.

2.4. Retrieval strategy is used to retrieve L2 linguistic resource and background knowledge to attempt the test tasks. It includes recombining the target language knowledge, transferring L1 linguistic knowledge, translating L1 to L2, recalling, \*remembering L2 linguistic knowledge and \*using pictures to recall L2

linguistic knowledge. However, none of the students reported using any of recombining the target language knowledge strategies. Translating L1 to L2 was reported by the majority of the students in all the three task types. For instance,

*Transferring L1 linguistic knowledge*

L3.1:[I] sometimes used Thai grammar to construct sentences.

*Translating L1 to L2*

H2.3: .. and [I] translated Thai into English for some tasks.

H3.3: [There] might sometimes be Thai into English translation such as the sentences related to incomplete tour program that I need to apologize the tourists. I use what I learned from the class.

*Recalling*

H1.2: [I] reviewed and reread the answer to recall L2 structures.

*Remembering L2 linguistic knowledge*

H1.1: [I] memorized the grammatical structures; for example, “It was built in the reign...”

H1.3: [I] emphasized on memorizing word orders such as “Next” that was used to explain the next place.

H2.1: [I] thought of important L2 structures. For example, [I] wanted to warn the tourists to keep their wallet [so I] needed to translated some parts in Thai and translated back into L2 as “Please [be] careful for your wallet”.

*Using pictures to recall L2 linguistic knowledge*

H2.2: [I] ..... also used the pictures. For example, when [I] see people with [an] undershirt picture [I] must warn the tourist to dress politely because it was the [an] important ritual site.

**3. Metacognitive strategy** deals with the conscious investigation of the test taking process to identify, organize, evaluate and plan the effective ways in attempting the test tasks. It is made up of three sub-categories: goal setting, assessment and planning.

3.1. Goal setting strategy involves identifying the tasks and deciding what to do. In this strategy, the students identified the different parts of the tasks before organizing their thoughts to respond to the tasks. This strategy was primarily reported by the high proficiency group across the three task types. The students identified the information in the tasks only in Task type one and three whereas identifying the prompt was employed in all three task types. The students identified the situations in the tasks and organized their thoughts only in Task type three. Following are some examples.

*Identifying the information in the tasks and deciding what to do*

H1.2: For example, with the picture or tour program [I] needed to understand first and decided on what to do.

H3.2: [I] listened and [tried to] understand [the conversation]. For example [in the] complaint [situation I] needed to apologize first and then tried to find the supporting reason to respond [to the task].

L3.1: [I] found out on what the tourists wanted to prepare the answer.

*Identifying the prompt and organizing the information*

H1.3: [I] found out what was the requirement from the prompt and selected the information [and] arranged them in my own words.

H2.2: [I] focused on the prompt and prepared the information to respond to the prompt. Then, [I] selected the information that was related to each task.

H3.1: [I] focused on what the requirement from the prompt was, thought about what the tourists wanted, tried to find the vocabulary and put them into sentences.

*Identifying the situations and deciding what to do*

H3.3: [I] primarily focused on the situation to decide how to respond.

L3.2: [I] listened and figured out the situation to answer the question.

3.2. Assessment strategy involves assessing what is needed, what one has to work with, and how well one has done. The majority of the students assessed how well they have done and what was needed to respond to the three task types whereas only the student from High 3 group reported on what to work with. Examples are given below.

*Assessing what is needed*

H2.1: [I] thought of technical terms and some sentence structures to use with like the prohibition.

L1.2: First, [I] assessed on what the prompts asked before attempting it to answer the question.

L2.1: [I] tried to think about the information to respond to the task.

L3.2: [I] listened to the conversation in the task and thought about the answer for each task.

*Assessing how well they have done*

H1.1: [I] thought about how well the tourists could understand the answer.

L1.1: [I] thought about how well I did for each task.

H2.2: After [I] finished the task, [I] assessed how well [I] had done, but [I] could not change anything. [I was allowed to] record only once.

L3.1: [I] thought about how well I did whether I used the right grammar [and] vocabulary with the reason. I wondered how well the tourists could understand my answers.



*Thinking of what they have to work with*

H3.2: [I] figured out what the tourist talked. We needed to find the reason to respond right away. There was no preparation time such as the question about the tourists' attraction.

3.3. Planning strategy deals with deciding how to use the existing linguistic knowledge and background knowledge. The majority of the students reported using planning strategy in the three task types. The students decided to use current linguistic knowledge to attempt the task by organizing the existing grammatical sentence structures and vocabulary. For the background knowledge, they used the tourism knowledge that was relevant to the prompt and situations. Examples of the two categories are listed below.

*Organizing grammatical sentence structures and background knowledge related to prompts*

H1.1: [I] thought of how to use the tourism information to respond [to] the prompt. And which structures should be used?

H1.2: Yes, I needed to use tourism knowledge to respond to the task. But, if I struggled, I put my opinions especially for the recommendation for the tourists. Then, I think of grammatical structures.

H2.2: [I] used, for example, with the bus picture. [I] thought of how to tell the tourists. [I] need to have tourism knowledge to back up my explanation that they could take the bus. They did have to walk and I used "You should take transfer BTS". [I] think of structures of the sentences.

L2.2: [I] tried to think of how to use the tourism information to warn the tourists. [I] focused on Thai cultural information on prohibition. Then, [I] made sentences.

This includes deciding on how to use grammatical sentence structures and background knowledge related to each situation. It was only found in Task type three.

H3.1: [I] found the sentence that [is] related to the situation and used the information from the class to respond [to] the prompt by arranging the sentences.

H3.2: [I] figured out what reason to be used to answer the question with the focus on general tourism knowledge to resolve the situation. After that [I] made sentences.

L3.2: [I] selected the tourism information to answer the particular task with appropriate grammatical structures.

L3.3: [I] practiced to be [a] tour guide to resolve the situation. So, [I] must know how to choose the information and speak English sentences that were appropriate with a particular situation. [I] focused on the tourists' requirement.

*Organizing vocabulary and background knowledge related to prompts*

H1.3: [I] selected the most relevant [part] to the prompts; for example, What was the architectural structure of the Summer Palace? [I] needed to think of the vocabulary.

H2.1: [I] used lots of [background knowledge] especially Cultural Knowledge particularly for the technical terms. For example "Please be careful" and [I] found the reasons to support the answer related to Thai culture.

H2.3: Sometimes, [I] thought of the technical terms and specific reasons on Thai culture to answer the prompt.

In summary, the high proficiency group reported the total of 13 sub-strategies while 11 of them were reported by the low proficiency group. The high proficiency students reported almost half of the total number of strategies, and the most reported strategies were Cognitive, Metacognitive and Communication respectively. The four most reported sub-categories were Comprehending, Storing memory, Retrieval and Planning. The first three sub-categories were from Cognitive and the last one was

from Metacognitive strategy. The result also showed the highest frequencies of the reported strategies in Task type three whereas the similar percentage of strategies was found in Task type one and two. However, the difference was not so obvious.

#### **4.4 Summary**

To summarize, this chapter reports on the results of the findings. For the first research question, the results from the two-way ANOVA showed that only the proficiency levels had a significant main effect on LSP speaking performances. The high proficiency students posed higher total scores than the low proficiency students across the three task types. However, their performances in the three task types were not significantly different, indicating that the high proficiency group's performances were stably high while those of the low proficiency group were constantly low across the three different task types. More specifically, the effect of task types was found in relation to the proficiency levels. The task types affected more speaking components in the low proficiency group than those of the high proficiency group. The low proficiency group had significantly different performances in language functions, fluency and content knowledge among the three task types while such difference occurred in vocabulary and content knowledge performances in the high proficiency group. In addition, content analysis showed both similarities and differences in each LSP component of the two proficiency groups, and some were associated with a particular task type and proficiency levels. With regards to the second research question, the two proficiency groups' views towards the WBST-EFT were not significantly different and they also had positive views on all aspects of the test. For the last research question, it was found that students with different proficiency levels employed different types and frequency of the speaking test taking strategies. The high proficiency students reported a wider type of strategies, but the difference was not notable. They reported higher frequencies of test taking strategies than the low proficiency students across the three task types. The findings also showed the difference in the total reported strategies from the two proficiency groups among the three task types, but such difference was not so obvious. In the next chapter, the summary of the study, discussions on the findings and recommendations for future studies will be given.

## CHAPTER V

### DISCUSSIONS AND CONCLUSIONS

#### 5.1 Summary of the study

This study aimed to study both the main and interaction effects between the task types of the Web-based Speaking Test in English for Tourism (WBST-EFT) and English proficiency levels on students' speaking performances, examine the students' attitudes towards this LSP online speaking test, and investigate and compare the types and frequency of strategies in doing the test. The total of 120 third year university students participated in the study in the second semester of the academic year 2010. The research instruments were the needs analysis questionnaire, the WBST-EFT and the rating scale, the attitudes towards the WBST-EFT online questionnaire and the speaking test taking strategies interview questions. The English for Tourism speaking performances were investigated on the knowledge of pronunciation, vocabulary, grammar, language functions, cohesion, fluency and content knowledge. The test comprised three task types: presenting tourism-related information, giving polite suggestions to the tourists and responding to tourists' enquiries and complaints. The data were quantitatively and qualitatively analyzed. The two-way ANOVA, independent samples t-test and descriptive statistics (means, standard deviations, percentage and frequency) were the quantitative approach, while the qualitative approach included content analysis of the test responses, attitudes towards the test, and verbal reports on the test taking strategies. The limitations of this study related to the small number of the sample size and the limited coverage of the task types used by tour guide professionals. The task types were selected in relation to the English for Tourism II course final achievement test content; hence, the test might not cover all the actual tasks in the tour guide context. The findings of the present study are as follows.

First, the findings from the total mean scores of the two proficiency groups showed that only the proficiency levels had a significant main effect at .05 level on the speaking performances and the effect size was relatively small. There was neither a significant main effect from the task types nor interaction effect between the three

task types and the two proficiency levels. In other words, only the proficiency levels had a considerable impact on the students' LSP speaking performances. The speaking performances from the two proficiency groups were constant across the three task types. With regards to the means of the individual speaking component, there were significant main effects from both the proficiency levels and task types on vocabulary, language functions and content knowledge at .05 level. The largest effect size of both proficiency levels and task types was in content knowledge. This indicated that the mean scores of these components from the two proficiency groups were different among the three task types and the most difference in the mean scores was found in the content knowledge. More specifically, a Scheffé post-hoc test revealed that the significant mean difference of vocabulary was found between Task type one and two. For language functions and content knowledge, the two proficiency groups performed differently and significantly between Task type one and two, and Task type one and three respectively. Considering the total scores in each proficiency level, a significant difference was found in vocabulary and content knowledge for the high proficiency students with the small effect size. However, the effect size of the content knowledge was more than twice as many values as the vocabulary. More specifically, results from a Scheffé post-hoc test showed that a significant mean difference in vocabulary was only between Task type one and two, while the mean difference in content knowledge was between Task type one and two, and Task type one and three. For the low proficiency students, the significant mean difference was in language functions, fluency and content knowledge with the small effect size. Similar to the high proficiency group, the effect size of content knowledge was almost four times higher than that of the language functions and fluency. A Scheffé post-hoc test displayed a significant mean difference between Task type one and two, and Task type one and three for language functions and content knowledge. While the sole significant mean difference was found between Task type one and three in the fluency component. All the significant differences were at .05 level. Additionally, the qualitative content analysis revealed similarities, differences, typical errors and prominent errors in particular task types in each LSP speaking component and some were associated with particular task types and proficiency group.

Second, the two proficiency groups' views towards the WBST-EFT did not differ on the four aspects: overall usefulness, appropriateness of time for preparation and response formulation, task difficulty, and the interface design. Their attitudes were found to be positive. The content analysis also showed that the two proficiency groups liked the multimedia and relaxing test taking procedures which were regarded as the strengths of this LSP online test. The correspondence between the situations on the test tasks and the real world tasks were considered as the strength of this instrument. However, the technical problem, especially on the sound recording system was considered the weakness of this technology-based test.

Finally, findings on the total reported speaking test taking strategies indicated that the high proficiency group reported a total of 13 sub-categories and 11 of them were from the low proficiency group, indicating that they employed almost similar types of strategies to attempt the test tasks. The high proficiency group used almost twice as many strategies as the low proficiency group across the three task types. Cognitive was the most frequently reported strategy, followed by Metacognitive and Communication respectively. The most differently reported strategy between the two proficiency groups was in Cognitive which was employed by the high proficiency group at almost twice as frequently as the low proficiency group. Additionally, Comprehending, Storing memory, Retrieval and Planning were the most frequently used sub-categories by the two proficiency groups. The first three ranks were from Cognitive whereas the last one was from Metacognitive strategy. Considering the task types, the two proficiency groups employed the highest frequencies of strategies in Task type three while similar frequencies of strategies were reported in Task type one and two. However, the difference was not notable. Cognitive was the most frequently reported strategy in Task type three, and it was similarly used in Task type one and two. It was followed by Metacognitive that was frequently reported in Task type three. Communication was the least reported strategies across the three task types.

## 5.2 Discussions

### 5.2.1 Effects of the WBST-EFT task types and proficiency levels on speaking performances

- *Main and interaction effects between the task types and proficiency levels on the total scores of the speaking performances*

The results from the two-way ANOVA showed a significant main effect of the proficiency levels on the total scores of the speaking performances at .05 level with a relatively small effect size. Although the value of the effect size was not very large, it was the only significant difference found in the present study. The results, hence, indicated a considerable impact of the proficiency levels on the speaking performances with the high proficiency group performing better than the low proficiency group across the three task types. In addition, the total mean scores from the three task types were not significantly different, thus, indicating no significant main effect of the task types on the speaking performances. Likewise, there was no significant interaction effect between the two proficiency levels and the three task types on the total means of this LSP test performances. In other words, the two proficiency groups' total scores were stable across the three task types.

- *Main and interaction effects between the task types and proficiency levels on the individual speaking component*

The comparison of the individual speaking component means showed significant main effects from both the proficiency levels and task types on vocabulary, language functions and content knowledge. More specifically, a Scheffé post-hoc test revealed the difference among these components. In vocabulary, a significant difference was found between the means of Task type one and Task type two. It may be due to the types and number of different technical terms involved in the task requirement. In Task type one, a particular field of technical terms related to palace and temple architectural structures, Thai history and Buddhism were the primary requirement in this task type. In contrast, Task type two elicited the students' vocabulary knowledge in both Thai etiquettes technical terms and generic terms

whereas in Task type three, only generic terms were required. For language functions and content knowledge, the differences were between the means of Task type one and two, and Task type one and three. It may be that both Task type two and three shared similar features on the slight integration of these two components in the students' performances. In Task type two and three, the two proficiency groups employed similar manipulative function in these task types. For content knowledge, these two proficiency groups employed the tourism content knowledge in Task Four (Task type two) to provide polite suggestions to the tourists at the attraction. In Task type three, the content knowledge was about dealing with tourists' enquiries and complaints. These types of content knowledge in Task Four and in Task type three may be similar in that they were more general than the content knowledge in the previous test tasks, which involved specific types of content knowledge in Thai history, arts, culture, and etiquettes.

- *Main effect of the task types in each proficiency level on the individual speaking component*

In addition, the results from the one-way ANOVA reconfirmed the main effect of the task types in relation to the proficiency levels on some of the speaking components. The findings indicated the different impact of the task types in each proficiency level. More specifically, the task types affected more speaking components with a higher value in the low proficiency group than those of the high proficiency group. In the high proficiency group, the significant mean difference was in vocabulary and content knowledge. A Scheffé post-hoc test showed that the high proficiency group had different vocabulary scores between Task type one and two. Their content knowledge means differed between Task type one and two, and Task type one and three. For the low proficiency group, significant mean differences were found in language functions, fluency and content knowledge. A Scheffé post-hoc test showed that their language functions and content knowledge means were significantly different between Task type one and two, and Task type one and three. Their fluency means differed between Task type one and three. The difference of task type effects in the proficiency levels may be from the proficiency levels in the target language. It is noticeable that the high proficiency group was able to maintain constantly high scores



of most of their LSP individual component's performances across the different task types. Because of the low proficiency level in the target language, more LSP components' performances of the low proficiency group differed among the task types than those of the high proficiency group as can be seen from the means differences in three components of the low proficiency group. Moreover, the additional significant difference in the low proficiency group's performances, particularly in the fluency component that differed between Task type one and three may be due to the amount of information required from the test task. It is obvious that Task type two and three require relatively similar amount of information in that they did not elicit much information from the students as much as Task type one. In Task type two, the students were asked to provide information on do's and don'ts at the ritual and tourist attractions, while information on the problem-solving in each tourism situation was in Task type three. In contrast, Task type one requires the highest amount of information among the three task types since it aims to elicit the students' ability in explaining Thai history and architectural structure in the full details.

The findings of the task types effect on some of the LSP speaking components are supported by some previous studies (Teng, 2008 and Kim, 2009). Teng (2008) investigated the effect of three task types: answering question, describing pictures and presentation on EFL speaking performance regarding accuracy, complexity and fluency. This author found a particular task type, answering questions, had an effect on complexity and fluency of participants' performances. Similarly, Kim (2009) investigated the effect of context and task types, independent and integrated skill tasks, on second language speaking ability. This author also indicated a small effect of task types on some speaking components, sociolinguistic competence and task completion, on L2 speaking ability.

- *The stable LSP speaking performances across different task types*

The quantitative findings on "no effects of task types" on speaking performances disagree with the previous studies (Turner & Upshur, 1995, Lumley & O'Sullivan, 2005 and Kim, 2009). There are many possible reasons for the indifferent

effects on test performances across task types which are all associated with proficiency levels.

First, proficiency in language knowledge may be the prominent factor influencing language performances in all different task types. Since all task types in WBST-EFT are context specific, which reflected the nature of LSP test, it may be possible that the students mainly rely on their language knowledge to respond to these test tasks. To be precise, the high proficiency students were able to maintain their high speaking scores across different task types while the low proficiency students' performances scores were also low in all task types. The speaking scores in this study were on linguistic and content specific knowledge. The results from the two-way ANOVA indicated no significant task types effects on the majority of speaking components, particularly on pronunciation, grammar, cohesion and fluency. In other words, the mean scores from the two proficiency groups of these speaking components were relatively stable across the three task types.

Statistical evidence indicated that the two proficiency groups made use of their language knowledge primarily associated with their proficiency levels, negligible of the different task types. More specifically, as the students have higher proficiency levels, they are more capable to manipulate their language competency to attempt different task types than those with the low language ability. The results from the comparison of mean difference in each proficiency group indicated that the high proficiency group's speaking components means were more relatively stable across task types than those of the low proficiency group. The mean difference occurred only in vocabulary and content knowledge, thus, reflecting the task type effects on these two components. It is noticeable that the high proficiency group's mean scores on pronunciation, grammar, language functions, cohesion and fluency were constantly high across all task types, thus, reflecting that they were effectively capable of manipulating these language components in their speech performances. Due to low proficiency in language knowledge, the majority of their speaking components' means were noticeably low and relatively constant across the three task types. However, task type effects were more evident in the speaking components of the low proficiency group than those of the high proficiency group. The means difference was

found in language functions, fluency and content knowledge. These language components' means were found to be different among the three task types. Moreover, the effect size of task types on speaking components in the high proficiency group was smaller than that of the low proficiency group, indicating less means difference among the task types. A lack of effect of the task types on the total speaking performances was also found in the study by Teng (2008). The author indicated no significant difference in the total means of speaking performances across the three task types. Similarly, Fulcher and Marquez-Reiter (2003) infer that language ability had greater effect on the scores than on the task conditions.

The second reason may be due to cognitive processing in responding to the stimuli in the prompt associated with proficiency levels (Anderson, 1983 cited in Van Moere, 2006). In this study, all the three task types required the students to process and transform cognitively complex stimuli which require multi skills. Task type one requires the students to read and respond to motion visuals and written texts. Task type two requires the students to listen and respond to the motion pictures and Task type three similarly requires the students to watch and listen to video clips. It is clearly seen that all the test tasks were integrated-skill oriented with complex stimuli. The students had to similarly manipulate their cognitive process to interact with the complex input and formulate their responses across the task types. Cognitive processing in this study is part of the speaking test taking strategies, and the findings showed no distinct difference in the frequency and types of strategies used across tasks. More specifically, similar frequency of Cognitive strategies was found in Task type one and two, whereas almost similar frequency of strategies was found in Task type three. Owing to the similarity in frequency and type of Cognitive strategies, it possibly resulted in the stable test performances across the different task types. However, a prominent difference was associated with proficiency levels. The high proficiency students reported almost twice as many Cognitive strategies as the low proficiency group. They possibly benefit from their cognitive process in interacting with prompts and formulate their responses by using less time in these processes, but could produce more amount of speech than the low proficiency students. The amount of speech production was part of the marking criteria in this study, and the high

proficiency students' performances in this study posed this criterion across the different task types. This cognitive advantage is based on cognitive development proposed by Anderson (1993 cited in Van Moere, 2006) in that the high ability test takers may be more advantageous since they spend less time on processing the test input to formulate their speech than the low ability test takers; indicating that the difference in proficiency levels plays an essential role in input processing and response formulation in the test.

Finally, the stable performances across the different task types may be explained by the way the students used their strategic competence which is assumed to be operative in all of the communicative situations (Douglas, 2000:38); and in this study it is part of Metacognitive strategies. The findings showed almost similar frequency and types of Metacognitive strategies in Task type one and two from the two proficiency groups. However, there was one difference in the frequency of strategy used in Task type three which was primarily derived from the low proficiency group. The low proficiency group reported the highest frequency of Metacognitive strategies in the last task type with almost similar percentages of the three sub-strategies, Goal setting, Assessment and Planning. In contrast, the high proficiency group reported similar type of sub-strategies and constantly high frequency with no prominent difference across the three task types. It can be seen that there was a difference in the low proficiency group only in Task type three which might be from greater processing demand in attempting the last task type than the previous two task types due to the type and amount of input in the prompt. Comparing to the previous task types, Task type three requires more frequency of higher order thinking than the previous two task types. In Task type three, the students were provided with different situational video clips without any written texts, and they were also asked to listen and immediately respond to the prompts while the previous two task types came with written texts and pictures in the prompts. This interactive media could assess complex Metacognitive ability of the students and facilitate more realistic tasks (Hamilton, Klein & Lorie, 2000 and Garcia Larboda, 2007a). From the present study's findings, the two proficiency groups constantly employed almost

similar types and frequency of strategic competence in attempting test tasks which may, therefore, lead to stable performances across the task types.

- *The distinct features of some speaking components in relation to task types*

The qualitative content analysis on the test responses from the two proficiency groups showed distinct features of some speaking components associated with task types, specifically on the use of certain types of vocabulary, tenses, grammatical structures, language functions, cohesive markers and content knowledge. These distinctive features in the responses of the two proficiency groups may be from the specific purpose input in the test content and task requirement which reflected the characteristics of an LSP test (Douglas, 2000).

Concerning the use of vocabulary and corresponding to the statistical difference, a particular type of terms associated with the task types was prominent. Tourism technical terms, specifically Thai history and architectural structures were primarily found in Task type one whereas Thai etiquettes terms were mainly used in Task type two, especially in Task Three. For grammar, the frequency counts showed that present simple and future 'will' tenses were found across three task types while past simple tense, particularly the passive voice was mainly used in Task type one to explain about the period of construction of the attractions. Simple and compound constructions were primarily used in Task type one and three. These sentence types were used in Task type one to introduce the attraction, and they were employed in Task type three to respond to tourists' enquiries and complaints. Complex sentences were mainly used in Task type one and two to point out the importance of the attractions in Task type one and to provide the reasons for giving suggestions to tourists in Task type two.

The analysis showed particular structures in a certain task type. In Task type one, the use of "*Pronoun+ V to be + Article + Noun*" was to introduce the attraction. Additionally, "*Pronoun+ V to be + Past participle+ Complement*", "*Noun+ V to be + Quantifier+ Adj+ Conjunction+ Complement*", and "*Pronoun+ V to be + Quantifier+ Adj + Noun+ To + V+ Complement*" were employed to explain the information about the attractions. "*Cohesive Marker+ Prep+ Time+ Future 'will'+ Complement*" was used to

explain the tour itinerary. In Task type two, the structures “*Adv ‘Please’+V+ Conjunction+ Complement*”, and “*Pronoun+Modal + ‘should’+V+ Conjunction + Complement*” were mainly used for giving suggestions on what the tourists should do in Thai culture and at the attraction while the use of “*Adv ‘Please’+ Auxiliary not+Conjunction+Complement*”, and “*Pronoun+Modal ‘should not’+V+Conjunction + Complement*” were for warning the tourists about the things that they should not do. In Task type three, the structures of “*Auxiliary not+ Verb+Sir/Madam. Pronoun+Future ‘will’+Complement*”, “*Pronoun+Vtobe+Adj+Sir/Madam.Pronoun+ Future ‘will’+Complement*”, and “*Exclamation ‘OK’+Sir/Madam.Pronoun+Future ‘will’+Complement*” were used in both dealing with tourists’ enquiries and complaints tasks.

For language functions, a particular feature was notable in a certain task type. In Task type one, the analysis showed that the ideational function was used in explaining the temples and palaces. Manipulative function in giving polite suggestions about Thai etiquettes and tourist attractions was found in Task type two. This language function was also found in Task type three in providing solutions to the tourists. On the use of cohesion, the prominent feature was noted in some task types. The frequency counts showed that time sequence cohesive markers were mainly used in Task type one, specifically in Task Two. The connector ‘because’ was mainly used in Task type two while no specific feature was found in Task type three. Additionally, a particular type of content knowledge associated with the task types was salient. In Task type one, content knowledge was noted in Thai architectural structures, Thai arts and history and Buddhism. In Task type two, Thai cultural etiquettes knowledge was noted, as well as in Task Three, while content knowledge on dealing with tourists’ enquiries and complaints in tourism context was prominent in Task type three.

- *The relationship between grammatical knowledge and the proficiency levels in the LSP online speaking test*

In addition, the analysis showed the relationship between grammatical knowledge and the proficiency levels in the LSP online speaking test. Considering quantitative results, the effect size value of proficiency levels in grammar was the

second largest with no task types impact. This finding supports the strong relationship between the proficiency levels and grammatical knowledge. The possible reason with no tasks effects on grammar may be that the high proficiency students effectively and accurately exploit their target language grammatical knowledge, particularly on the use of sophisticated constructions across different task types. Proficiency in grammar is also referred to as proficiency levels in the previous studies of the LSP tests (Clapham, 1996 cited in Douglas, 2000).

- *The prominent features of the LSP construct found in the WBST-EFT*

Interestingly, the content knowledge component contained the highest effect size from both the proficiency levels and task types among seven LSP speaking components. More specifically, the mean difference of both the high and low proficiency groups also indicated the highest effect size value in content knowledge component. From the statistical evidence, task types effect is notable on content knowledge. The findings thus reconfirm Douglas's (2000 & 2001) theory of the essence of the LSP construct in that the field specific background knowledge is manipulated in relation to the test task characteristics to interpret the communicative situations and formulate the responses. The impact of each task type on the content knowledge may be due to the characteristics of specific purpose input, specifically in the test content contextualized cues (Douglas, 2000:44) and task requirement. These LSP task features in each task type affected some LSP components of the test performances as shown in this study. In Task type one, the contents are related to the target language used to describe tourism places, particularly in Thai palaces and temple settings. The contextualized cues that marked the distinctive features of the LSP tasks were on particular settings and technical terms in Thai history and architectures. For task requirement, this task type aimed to elicit the students' speaking ability in ideational function, specifically explaining famous Thai palaces and temples in the central region; thus, the content knowledge in this task was Thai historical background and architectural structure. Albeit some similar features in the settings in the prompts, Task type two's target language use was on giving polite suggestions to the tourists on do's and don'ts in Thai culture and at the crowded tourism site. This task aimed to elicit manipulative function and the content

knowledge related to both Thai etiquettes in the temples and Thai manners in the tourism context. In Task type three, the settings were changed in relation to the situations at the hotel, on the bus and at the attractions. The content knowledge was on dealing with tourists' enquiries and complaints. In addition, these findings answer the critical issue on the effect of background knowledge and proficiency levels on LSP test performances. Proficiency levels had great impact on LSP test performances in speaking skill since the statistical evidence indicated that the high proficiency group posed almost twice as many content knowledge means as the low proficiency group. In other words, as the students are more proficient in their target language, it is more possible for them to effectively manipulate the background knowledge ability. This finding corresponds with the finding of Clapham (1996 cited in Douglas, 2000) in IELTS reading test in which high proficient test takers benefited from their subject areas content knowledge while the lower level test takers did not gain any benefits from their specific area knowledge.

- *The content and construct of the LSP online speaking test in English for Tourism*

Finally, the quantitative and qualitative findings in this study revealed the content and construct of the LSP speaking test in tourism context with the web technology-based assessment. Significant mean differences revealed that LSP ability contains specific vocabulary, language functions, and most importantly specific content knowledge. Content analysis also indicated specific features of vocabulary, grammatical structures, language functions, cohesion and content knowledge in relation to task types. These distinctive features associated with the specific purpose input in task characteristics resulted from target language use analysis and subject specialists' views (Douglas, 2001). Consequently, the test task content and methods closely corresponded to the real world tasks and indicated the authenticity and validity of the WBST-EFT as the test performances can be inferred to the real world performances, particularly in Thai tourism context.



### **5.2.2 Students' attitudes towards the WBST-EFT**

The findings from the independent samples t-test revealed that the two proficiency groups' attitudes towards this online speaking test were not significantly different in general. The high proficiency group had slightly higher total means than the low proficiency group, and the two proficiency groups had high mean scores in the four aspects. In other words, the two proficiency groups similarly agreed that the WBST-EFT was useful, had appropriate preparation and response formation time, appropriate interface design, and the test tasks were not too difficult. From the statistical evidence, it can be claimed that their views towards the online test were similar and positive.

With regards to the positive attitudes of the two proficiency groups, it may be attributed to the appropriate use of the interface design, particularly in the test tasks presentation. The integration of multimedia, especially the motion pictures and video clips in the test tasks, could facilitate the students to recall the memory and memorize all information, particularly in Task type one in which excessive amount of information was required. This included the page layout that contained appropriate details, clear titles and easy reading texts. Furthermore, the correspondence of test tasks to the real world tasks may have also supported the positive attitudes that the students can possibly use their knowledge about the real world situations. The relaxation in test taking procedures may also support the positive views to this online speaking test which was a semi-direct speaking test that excluded the interlocutor. The students, therefore, could take the test without supervision and felt at ease in this friendly atmosphere. The preparation and response formulation time may have contributed to the positive opinions towards the WBST-EFT, since the students had enough time to gather all the required information and formulate their responses.

The findings of the positive attitudes are consistent with previous studies (Kenyon & Malabonga, 2001 and Warschauer, 1996). Kenyon and Malabonga (2001) compared the test takers' attitudes between the tape-based Simulated Oral Proficiency Interview (SOPI) and Computerized Oral Proficiency Instrument (COPI). The authors found that the test takers preferred COPI and thought that it was more advantageous

than the previous version, specifically on the issues of fairness of the speaking situations, task difficulty and test taking nervousness. Similarly, the finding of Warschauer (1996) reconfirmed the positive attitudes on the Computer-based Test (CBT) comparing to the direct face-to-face speaking test. The participants felt at ease and comfortable via the CBT and thought that it facilitated their thinking ability. They liked the task delivery system which was part of the interface design in the present study.

Although no statistical evidence was found, the content analysis of the open-ended parts showed the students' negative attitudes towards the recording system of the WBST-EFT. They stated that saving and sending files online were too demanding for them. This negative attitude may be from the administration time constraint and a number of requirements. At the beginning, the students had to orientate themselves to the test taking procedures. After they had completed the test, they were required to save the files and submit them online via Moodle platform so that the files would be in the online database. Then, they had to complete the online questionnaire. All these procedures must be finished within 45 minutes.

It is interesting that a number of scholars mention about the advantages of the technology-integrated speaking, especially the capability of eliciting and assessing more complex language production than the traditional face-to-face speaking test (Warschauer, 1996, Norris, 2001 and Garcia Larboda, 2007a). The appropriate incorporation of the interactive input, particularly the multimedia in the test tasks, therefore, requires validation procedures (Chapelle, Jamieson & Hegelheimer, 2003) and the students' reflections in this study provided insightful information for such procedure.

### **5.2.3 Speaking test taking strategies used by the high and low proficiency students in doing the WBST-EFT**

The frequency and percentage from the verbal reports analysis showed a difference in the number and types of the reported strategies between the two proficiency groups. Additionally, the findings indicated no distinct difference in the

number and types of strategies used by the two proficiency groups across the three task types.

From the total number of the reported strategies, the high proficiency group reported 13 sub-categories whereas the low proficiency group used 11 sub-strategies, showing no obvious difference in the types of reported strategies between the two proficiency groups. However, the high proficiency group reported almost half of the total strategies. Cognitive was the most frequently reported strategy by the two proficiency groups. It was also the most differently reported strategy between the two proficiency groups. It was followed by Metacognitive and Communication respectively. The two proficiency groups reported almost similar number of Communication strategies. Considering the sub-categories in the individual strategy, Comprehending, Storing memory, Retrieval and Planning were the most frequently reported sub-categories by the two proficiency groups. The first three ranks were from Cognitive whereas the last one was from Metacognitive strategy. Retrieval was the most differently reported sub-category between the two proficiency groups whereas both Comprehending and Paraphrase were equally reported by the two proficiency groups. The possible reasons that Cognitive was the most frequently reported strategy by the two proficiency groups may relate to the way they understand and respond to the test tasks. It can be seen from Comprehending strategy that it was the most reported sub-category by the two proficiency groups. The two proficiency groups also reported a similar amount of this strategy across the three task types. It may be that the students had to know what to do in the task in order to prepare the information and formulate their responses regardless of the proficiency levels. The results from the content analysis of the verbal reports also showed that the students employed this strategy to understand both the instruction and requirement of the tasks and the prompts. They mainly analyzed the tasks in order to prepare the information to respond to the tasks. Similarly, the two proficiency groups reported similar frequency of Paraphrase for the main purpose of continuing their speech. However, Retrieval was the most differently reported sub-category between the two proficiency groups in that the high proficiency group employed four times as many strategies as the low proficiency group. The possible reason may be that the high proficiency group had

higher proficiency level of the existing L2 linguistic resource and background knowledge than that of the low proficiency group; therefore, the high proficiency group was more likely to be able to use this strategy to attempt the test task.

With regards to sub-categories in the individual strategy, Metacognitive, the two proficiency groups reported using Planning, Assessment and Goal setting respectively. The most differently reported sub-category was Goal setting. The high proficiency group frequently used almost three times as many Goal setting strategies as the low proficiency group. In Communication, the two proficiency groups reported more than twice as many Achievement strategies as Avoidance. The most differently reported sub-category was Approximation which was solely found in the low proficiency group.

The findings on the relationships among the proficiency levels, frequency and types of reported strategies correspond with the previous studies (Song, 2005, Swain et al., 2009, Cabaysa and Baetiong, 2010 and Mendez Lopez, 2011). The study of Cabaysa and Baetiong (2010) found a significant difference in the frequency of strategies used by the students from different levels of speaking proficiency, particularly in the use of Metacognitive strategy. Similarly, the study of Song (2005) on the exploration of the language strategy used in Michigan English Language Assessment Battery (MELAB) and the relationship between the strategy used and learners' test performance indicated a linear relationship between the strategy used and language test performance. The choice of strategy used by a particular proficiency group was also found in the recent study of Mendez Lopez (2011) in that the high proficiency learners reported more complex speaking strategies than the lower level language learners. The study on the technology-based speaking test of Swain et al. (2009) on the speaking strategies used in the internet-based TOEFL reconfirmed the previous findings that a particular type of strategy was associated with the proficiency levels. Communication strategies were employed more by the undergraduate test takers than the graduate test takers while the graduate test takers outnumbered their counterparts in Cognitive and Affective strategies.

Concerning the task types and reported strategies, the differences in both types and frequency of strategies used across the three task types were not so obvious. However, Task type three contained the highest frequency of the three types of reported strategies from the two proficiency groups whereas the similar frequencies of strategies were reported in Task type one and two. It is notable that the highest frequency of strategies in Task type three was mainly derived from the low proficiency group, particularly on the use of Metacognitive and Cognitive strategies while the high proficiency students constantly reported high frequency and varied types of strategies across the three task types. Regarding the difference in frequency of the reported strategies between the two proficiency groups, Task type one contained the most different use of both Cognitive and Communication strategy while Metacognitive was differently employed in Task type two. Concerning the sub-categories in the individual strategy in relation to the task types, for Cognitive strategy, the highest frequency of reported sub-category by the two proficiency groups was Comprehending in Task type three whereas the most different use between the two proficiency groups was Retrieval in Task type two. For Metacognitive strategy, Planning and Assessment were the most frequently employed sub-categories in Task type three. The most differently reported sub-category between the two proficiency groups was Planning in Task type one and Goal setting in Task type two. For Communication, Paraphrase was the most reported sub-category from the two proficiency groups, particularly in Task type two whereas the most differently reported sub-category between the two proficiency groups was Restructuring in Task type three and that was only reported by the high proficiency group.

The possible reasons that Task type three contained the highest amount of reported strategy may be due to the task requirement on the cognitive loads of the students (Teng, 2008). In Task type three, the students were first required to listen to different scenarios; then, they had to figure out the situations, gather the required information to formulate their answer and immediately respond to the test task without any preparation time. It was clearly seen that there were relatively high cognitive-oriented actions to attempt this task type. Comparing the three task types, Task type three, therefore, required the highest cognitive load than the previous two

task types which had written texts and planning time. The most difference in the frequency of the reported strategies between the two proficiency groups in Task type one may be due to the task requirement, specifically in the content knowledge and excessive amount of information. In Task type one, the students were required to provide knowledge on Thai history, arts, architecture and Buddhism in relation to the prompts' topics. This task type required the highest amount of information among the three task types. It can be seen that the high proficiency group effectively employed relatively high amount of Cognitive strategy to comprehend the prompts and formulate their responses. The high proficiency group frequently reported this strategy almost three times as many strategies as the low proficiency group. For Communication strategy, the low proficiency group outnumbered the high proficiency group almost three times which was only found in this task type. It may be that the high proficiency group had higher target language linguistic repertoire than that of the low proficiency group. Therefore, the high proficiency group depended less on this strategy.

As stated by Cohen (1998), the effective use of a certain test taking strategy depends on a number of factors (cognitive style and flexibility, test taking strategies' background knowledge, time and a particular task) including language knowledge. The results of the reported strategies in the present study conform to that of the speaking components' findings in that the students from different proficiency levels reported different frequencies and types of the strategies used.

### **5.3 Conclusions**

This study aimed to investigate the effects of task types and proficiency levels on students' LSP speaking performances in English for Tourism course using the internet assessment. Students' views on the WBST-EFT were also explored. Similarities and differences in the types and frequency of speaking test taking strategies used by high and low proficiency students in doing the WBST-EFT were examined. The findings showed that proficiency levels were the main factor affecting LSP speaking test performances in all components including the technology-integrated test. Cognitive process and the use of strategic competence may have also

affected the students' speaking performances in attempting the test tasks and formulating their responses. These two variables were also associated with proficiency levels. However, the effect of task types existed in some of the speaking components, particularly in vocabulary, language functions and content knowledge. It is noticeable that content knowledge posed the highest effect size among all of the speaking components in both proficiency levels and task types, thus, representing the prominent feature of the LSP test and indicated the strong relationship between these two variables. Moreover, it was found that task types had greater impact on the low proficiency group than the high proficiency group, particularly on the content knowledge component. In addition, content analysis from the test responses revealed that proficiency levels had a great impact on the difference of the LSP speaking performances between the two proficiency groups in terms of accuracy, range, appropriateness and complexity. The analysis also indicated the distinctive features of some of the speaking components which were associated with the task types: specific types of words, grammatical structures, language functions, cohesive markers and content knowledge. With regards to the students' attitudes towards the WBST-EFT, the findings indicated no significant difference between the two proficiency groups and their views on this LSP online test were positive. However, the test could be improved with regards to its recording system and administration time. These reflections can provide insightful information for test improvement and be used as the validity evidence. Finally, frequency and types of speaking strategies used in attempting the WBST-EFT between the two proficiency groups also differed. The high proficiency groups reported almost half of the total strategies in the present study, and they used wider types of strategies than the low proficiency groups. However, the difference in types of reported strategies between the two proficiency groups was not notable. Cognitive was the most frequently reported strategy by the two proficiency groups and it was followed by Metacognitive and Communication strategies respectively. The findings indicated no prominent difference in both types and frequency of strategies across the three task types. From these statistical results and qualitative findings, it can be concluded that proficiency levels accounted for the variables in the LSP test performances using the internet to assess speaking ability.

## **5.4 Implications of the study**

### **5.4.1 Theoretical implications**

Theoretical implications from the present study are as follows. First, some useful insights concerning factors affecting LSP test performances were demonstrated in this study. The findings showed that proficiency levels were the main variable to the LSP speaking performances, particularly when all the tasks were context specific which indicated that students' performances were stable across the three task types. There were some task types effects on speaking components, particularly in knowledge of vocabulary, language functions and content knowledge. These components were the prominent features of the LSP test, especially the last component, content knowledge, which had the highest task types effects. Moreover, some components might not be susceptible to the change in task contexts, and thus, remain constant across tasks. These components were knowledge of pronunciation, grammar, cohesion and fluency. The constant performances of these components reflected that students heavily relied on their target language proficiency in attempting different task types. In addition, choice and frequency of strategies used in attempting the test tasks and formulating the responses particularly the use of Cognitive and Metacognitive strategies, may have affected the students' performances in this study and may have contributed to the constant performances across tasks. What is prominent in this study is that task types effects on the speaking components were associated with proficiency levels. That is, the task types affected more LSP components of the low proficiency group's performances than those of the high proficiency group. Second, the WBST-EFT underpins LSP test construct proposed by Douglas (2000) and Fulcher's (2003a, 2003b) speaking tests and the interface design. The findings provide insightful validity evidence about the construct of the LSP test, particularly in Thai tourism context on using the internet to assess speaking ability. This LSP construct should be applied to other LSP tests in speaking skill with the integration of technology. Additionally, findings from both students' attitudes and test taking strategies can offer insightful information regarding the test takers' characteristics (Bachman and Palmer, 1996) which is considered as an important variable to test performances. Students' views and their test taking procedures in this



LSP-oriented speaking test with internet-based assessment should, therefore, be used as part of the validity arguments in test development to substantively support the usefulness of the test. These qualitative findings can also be used as the triangulation of the data for research studies. Finally, the study yielded the evidence of types and frequency of strategies that supports the findings from previous studies of Song (2005) and Cabaysa and Baetiong (2010) who found that proficiency levels affected the frequency of the reported strategies. Similarly, a relationship between choice of strategies and proficiency levels was found in the studies of Swain et al. (2009) and Mendez Lopez (2011). As presented in this study, the findings yielded empirical evidence to the LSP testing field that proficiency levels could be the main factor to the variation in test performances across the differences in task types.

#### **5.4.2 Practical implications**

The present study offers a number of practical implications. First, the WBST-EFT was constructed with a particular theoretical framework and put through the validation procedures to ensure an acceptable standard. It incorporated both the concepts of LSP testing and WBT construction using technological advantages in test development, storage and administration, particularly with the cost-effectiveness in test construction and non-sophistication of computer knowledge requirement mentioned by Hamilton, Klein and Lorie, (2000), Roever (2001) and Garcia Laborda (2007b). Therefore, the WBST-EFT is an interesting tool and should be used with a large number of test takers to assess their English for Tourism knowledge. Despite the advantage of WBT in administration, the findings of this study can provide information for the university about the usefulness and limitations of the WBT for assessing students' speaking performances in a large class. In addition, qualitative findings on prominent features of speaking components associated with task types can provide information about the English for Tourism course content and instructional approach for teachers and for the course development at NRRU. Moreover, as a diagnostic test, errors in each speaking component can be used in the remedial course to improve students' LSP speaking skill. Finally, the findings on types of strategies reported by the two proficiency groups can be used in speaking training in English for Tourism II course. Students can be trained to use successful speaking strategies in

attempting the test to develop their speaking skill, particularly in the internet-based assessment.

### **5.5 Recommendations**

1. From the limitations of this study, the interactive part of speaking performances i.e. turn taking has not been investigated due to the semi-direct nature of the test. It will be more beneficial to the communicative testing field if this natural part of communication is included in the future research. Additionally, not all the language knowledge, task types and strategies were investigated. It will be more fruitful to include sociolinguistic knowledge in the test construct to see how this knowledge is operated in LSP speaking test in tourism. Moreover, more task types and test taking strategies, particularly Affective strategy, should be included in the next study. Additionally, the correlations among proficiency levels, types and frequency of strategies used in attempting the semi-direct LSP speaking test should also be investigated to yield more insights into LSP testing field. Furthermore, the sample group in this study was the students at NRRU and their majors were English, Business English, Japanese, French and Tourism Industry. Researchers should investigate other facets, particularly test takers' fields of study to find out whether their speaking test performances could be affected by proficiency levels related to other specific fields.

2. For test developers, there are some technological limitations that should be considered in the internet-based test construction. Storage of files is one of the drawbacks of Moodle that allows the maximum of 10 megabytes which could affect the clarity of the pictures and video clips on the tasks presentation. This storage includes audio files which could also affect the length of these files, particularly for the speaking test performances recorded in this format. In this study, additional storage was used which was Gmail that allowed extensiveness of file storage without any costs.

3. For language teachers, the findings showed that proficiency levels in the target language were the main and most prominent variable to language performances. The findings also showed that task types affected more LSP speaking components of the

low proficiency students than those of the high proficiency students. It is thus important for teachers to prepare students with sufficient mastery in the target language before attending the LSP course. In addition, LSP speaking skill also encompassed a number of knowledge such as pronunciation, vocabulary, grammar, language functions, cohesion, fluency and content knowledge, and these components should be appropriately integrated in the instructional approach and course content. Furthermore, teachers should also emphasize on the prominent features from the high proficiency students and errors in these speaking components to improve the students' LSP speaking ability. Since the findings showed that high proficiency students used high frequency and a wide range of speaking test taking strategies in attempting the test tasks, low proficiency students should be trained to effectively manipulate these strategies for successful speaking performances.

4. For replication purpose, there should be more subjects involved in the future study, particularly from the educational institutions that offer English for Tourism courses for the generalizability of the findings. Additionally, more administration time is suggested due to a number of requirements for the students to complete. They should be provided with more time to get oriented to the sample test and to respond to the WBST-EFT online questionnaire. Approximately 1 hour is sufficient to complete all the tasks.

5. Due to the importance of tourism enterprise that could generate a large amount of income to Thailand, a number of educational institutions has offered English for tourism courses. The WBST-EFT, which is a potential assessment tool that exploits the technological advantages with a particular theoretical framework and construct, could be modified and used with other tourism types e.g. Eco-tourism and Adventure tourism.

## REFERENCES

- Alderson, J. C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Alderson, J.C. (2009). Test review: Test of English as a Foreign Language TM: Internet-based test (TOEFL iBT). *Language Testing* 26 (4): 621-631.
- Anderson, J.R. (1983). *The Architects of Cognition*. Cambridge, MA: Harvard University Press.
- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.F. and Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Brown, A. (1995). The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing* 12 (1): 1-15.
- Brown, J.D. (1997). Computers in language testing: Present research and some future directions. *Language Learning and Technology* 1 (1): 44-59.
- Cabaysa, C. C. and Baetiong, L. R. (2010). Language learning strategies of students at different levels of speaking proficiency. *Education Quarterly* 68 (1): 16-35.
- Cambridge University ESOL Examinations. (2008). *International Legal English Certificate Handbook*. [Online] Available from: [http://www.legalenglishtest.org/downloads/ilec\\_handbook.pdf](http://www.legalenglishtest.org/downloads/ilec_handbook.pdf) [November 20, 2009]
- Cambridge University ESOL Examinations. (2009). *The Business Language Testing Service (BULATS) Test Specification: A Guide for Clients*. [Online] Available from: <http://www.bulats.org/docs/BULATS-Test-Specifications-a-Guide-for-Client-June09.pdf> [October 25, 2009]
- Campbell, D. T. and Stanley, J. C. (1966). *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally.

- Chapelle, C. A., Jamieson, J. and Hegelheimer, V. (2003). Validation of a web-based ESL test. *Language Testing* 20 (4): 409-439.
- Chapelle, C. A. and Douglas, D. (2006). *Assessing Language through Computer Technology*. Cambridge: Cambridge University Press.
- Clapham, C. (1996). *The Development of IELTS: a Study of the Effect of Background Knowledge on Reading Comprehension*. Cambridge: Cambridge University Press.
- Cohen, A. D. and Olshtain, E. (1993). The production of speech acts by EFL Learners. *TESOL Quarterly* 27 (1): 33-56.
- Cohen, A.D. (1998). *Strategies in Learning and Using Second Language*. New York: Addison Wesley Longman Limited.
- Cohen, A.D. (2006). The coming age of research on test-taking strategies. *Language Assessment Quarterly* 3 (4): 307-331.
- Cumming, A. (2001). ESL/EFL instructors' practices for writing assessment: specific purposes or general purposes? *Language Testing* 18 (2): 207-224.
- Cumming, A., Grant, L., Mulcahy-Ernt, P. and Powers, D. (2004). A teacher-verification study of speaking and writing prototype tasks for a new TOEFL. *Language Testing* 21 (2): 107-145.
- Davies, A. (2001). The logic of testing Languages for Specific Purposes. *Language Testing* 18 (2): 133-147.
- De Wet, F., Van der Walt, C. and Niesler, T. R. (2007). 'Automatic large-scale oral language proficiency assessment.' Proceeding of Interspeech 2007 Conference, Antwerp, Belgium (pp.218-221). [Online] Available from: [http://www.dsp.sun.ac.za/~trn/reports/dewet+vanderwalt+niesler\\_interspeech07.pdf](http://www.dsp.sun.ac.za/~trn/reports/dewet+vanderwalt+niesler_interspeech07.pdf) [September 10, 2008]

- Douglas, D. (2000). *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Douglas, D. (2001). Three problems in testing language for specific purposes: Authenticity, specificity and inseparability. In Elder, C., Brown, A., Grove, W., Hill, K., Iwashita, N., Lumley, T, McNamara, T. and O'Loughlin, K. (Eds.), *Studies in Language Testing: Experiencing with Uncertainty. Essays in Honor of Alan Davies* (pp. 45-52). Cambridge: Cambridge University Press.
- Edwards-Leis, C. (2007). 'Variations to stimulated recall protocols to enhance student reflection: I did, I saw, I remembered.' Paper presented at the AARE 2006 International Education Research Conference, Adelaide, Australia.
- Elder, C. (2001). Assessing the language proficiency of teachers: are there any teacher controls? *Language Testing* 18 (2): 149-170.
- Fraenken, J.R. and Wallen, N.E. (1993). *How to Design and Evaluate Research in Education* (2<sup>nd</sup> Ed.). New York: McGraw-Hill.
- Fulcher, G. (2003a). *Testing Second Language Speaking*. London: Longman/Pearson Education.
- Fulcher, G. (2003b). Interface Design in Computer-based Language Testing. *Language Testing* 20 (4): 384-408.
- Fulcher, G. and Marquez-Reiter, R. (2003). Task difficulty in speaking tests. *Language Testing* 20 (3): 321 - 344.
- Garcia Laborda, J. (2006). PLEVALEX: A new platform for oral testing in Spanish. *Eurocall Review* 9: 4-7.
- Garcia Laborda, J. (2007a). *From Fulcher to PLEVALEX: issues in interface design, Validity and reliability in internet based language testing*. CALL-EJ Online 9 (1).

- Garcia Laborda, J. (2007b). On the net: introducing standardized EFL/ESL exams. *Language Learning and Technology* 11 (2): 3-9.
- Gass, S.M. and Mackey, A. (2000). *Stimulated Recall Methodology and Second Language Research*. Mahwa, NJ: Lawrence Erlbaum Associates.
- Gass, S.M. and Mackey, A. (2007). *Data Elicitation for Second Language Research*. Mahwa, NJ: Lawrence Erlbaum Associates.
- Graduate College Program from Iowa State University of Science and Technology (ISU). (2009). *SPEAK/TEACH Test*. [Online] Available from: [http://www.grad-college.iastate.edu/speakteach/testing\\_information/overview.php](http://www.grad-college.iastate.edu/speakteach/testing_information/overview.php) [October 29, 2009]
- Gumperz, J.J. (1976). Language, communication and public negotiation. In Sanday, P.R. (Ed.). *Anthropology and the Public Interest* (pp. 273-292). New York: Academic Press.
- Halliday, M.A.K. (1973). *Explorations in the Functions of Language*. London: Edward Arnold.
- Halliday, M.A.K. (1976). *Learning How to Mean: Explorations in the Development of Language*. London: Edward Arnold.
- Hamilton, L.S., Klein, S.P., & Lorie, W. (2000). *Using Web-based Testing for Large-scale Assessment*. Santa Monica: RAND Corporation.
- Hong-Nam, K. and Leavell, A.G. (2006). Language learning strategy use of ESL students in an intensive English learning context. *System* 34: 399-415.
- Jamieson, J. (2005). Trends in computer-based second language assessment. *Annual Review of Applied Linguistics* 25 (2): 228-242.
- Kenyon, D., Malabonga, V., and Carpenter, H. (2001). Response to the Norris commentary. *Language Learning and Technology* 5 (2): 106-108.

- Kenyon, D., and Malabonga, V. (2001). Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments. *Language Learning and Technology* 5 (2): 60-83.
- Kim, H-J. (2009). *Investigating the Effects of Context and Task Type on Second Language Speaking Ability*. PhD thesis, Columbia University.
- Krekeler, C. (2006). Language for academic purposes (LSAP) testing: the effect of background knowledge revisited. *Language Testing* 23 (1): 99-130.
- Lam, W.Y.K. (2007). Tapping ESL learners' problems and strategies in oral communication tasks: Insights from stimulated recall. *Prospect* 22 (1): 56-71.
- Littlemore, J. (2003). The communicative effectiveness of different types of communication strategy. *System* 31: 331-347.
- Lumley, T. and O'Sullivan. (2005). The effect of test-taker gender, audience and topic on task performance in tape-mediated assessment of speaking. *Language Testing* 22 (4): 415-437.
- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal* 29 (6): 861-878.
- Ma, T. (2009). An empirical study on the comparison between strategies on oral English classes and writing English classes. *English Language Teaching* 2 (2): 39-45.
- Mackey, A., Gass, S., and McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition* 22 (4): 471-497.
- Malabonga, V., Kenyon, D., and Carpenter, H. (2005). Self-assessment, preparation and response time on a computerized oral proficiency test. *Language Learning and Technology* 22 (1): 59-92.



- Malone, M.E. (2007). Center for Applied linguistics. *Oral proficiency assessment: The use of technology in test development and rater training*. [Online] Available from: [http:// www.cal.org/resources/digest/oralprof.html](http://www.cal.org/resources/digest/oralprof.html) [October 29, 2009]
- Mendez Lopez, M. (2011). Speaking strategies used by BA ELT students in public universities in Mexico. *MEXTESOL Journal* 35 (1): 1-22.
- Nakatami, Y. (2006). Developing an oral communication strategy inventory. *The Modern Language Journal* 9: 151-168.
- Nakhon Ratchasima Rajabhat University. (2008). *Curriculum for Bachelor of Arts in English*. Nakhon Ratchasima: Nakhon Ratchasima Rajabhat University.
- Nakhon Ratchasima Rajabhat University. (2008). English for Tourism II Course Description for Bachelor of Arts in English (p 36). Nakhon Ratchasima: Nakhon Ratchasima Rajabhat University
- Norris, J.M. (2001). Concerns with computerized adaptive oral proficiency assessment: A commentary on “Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments” by Dorry Kenyon and Valerie Malabonga. *Language Learning and Technology* 5 (2): 99-105.
- Oxford, R. L. (Ed.). (1996). *Language Learning Strategies Around the World: Cross-Cultural Perspectives* (SLT&CC Technical Rep. No. 13). Mānoa, HI: University of Hawaii Press.
- Pearson Education. (2008). *The Versant Spanish™ Test: Test description & validation summary*. [Online] Available from: [http://www.versanttest.co.uk /pdf/ValidationReport.pdf](http://www.versanttest.co.uk/pdf/ValidationReport.pdf) [October 20, 2010]
- Polio, C., Gass, S.M, and Chapin, L. (2006). Using stimulated recall to investigate speaker perceptions in native-nonnative speaker interaction. *SSLA* 2 (8): 237-267.

- Poulisse, N. (1993). A theoretical account of lexical communication strategies. In: Schreuder, R. and Weltens, B. (Ed.), *The Bilingual Lexicon* (pp. 157-189). Amsterdam: John Benjamin Publishing Company.
- Roever, C. (2001). Web-based language testing. *Language Learning and Technology* 5 (2): 84-94.
- Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing* 11 (2): 99-123.
- Smith, B. (2003). The use of communication strategies in computer-mediated communication. *System* 31 29-53.
- Song, X. (2005). Language learner strategy use and English proficiency on the Michigan English Language Assessment Battery. *Spain Fellow Working Papers in Second or Foreign Language Assessment* 3: 1-26.
- Swain, M. and Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research* 37: 285-304.
- Swain, M, Huang, L., Barkaoui, K., Brooks, L. and Lapkin, S. (2009). Educational Testing Service (ETS). *The speaking section of the TOEFL iBT (SSTiBT): Test-takers' reported strategic behaviors (TOEFL iBT Research Report No.10)*. [Online] Available from: <http://www.ets.org/Media/Research/pdf/RR-09-30.pdf> [October 20, 2010]
- Teng, H-C. (2008). A study of task type for L2 speaking assessment. In M.Mantero, P. C. Miller, & J. L.Watzke (Eds.), *ISLS Readings in Language Studies* (pp. 433-446). St. Lois, MO: International Society for Language Studies.
- Thailand Tourist Arrivals. (2011). *Thailand Tourist Arrivals from 2007 to 2011 Per Quarter*. [Online] Available from: <http://www.thaiwebsite.com/tourism.asp> [March 2, 2012]

- Ting, S. and Phan, G.Y.L. (2008). Adjusting communication strategies to language proficiency. *Prospect* 23(1): 1-9.
- Tourism Authority of Thailand. (1979). *Tourism Authority of Thailand Act 1979*. [Online] Available from: [http://www.asianlii.org/th/legis/consol\\_act/taota1979337/](http://www.asianlii.org/th/legis/consol_act/taota1979337/) [October 20, 2010]
- Turner, C.E. and Upsher, J.A. (1995). Some effects of task types on the relation between communicative effectiveness and grammatical accuracy in intensive ESL classes. *TESL Canada Journal* 12 (2): 18-31.
- Van Moere, A. (2006). Validity evidence in a university group oral test. *Language Testing* 23 (4): 411-440.
- Wall, D. & Horak, T. (2006). *The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe: Phase 1, The baseline study*. TOEFL Monograph Series, MS- 34. Princeton, NJ: Educational Testing Service.
- Wall, D. & Horak, T. (2008a). *The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe: Phase 2, Coping with change*. TOEFL iBT Research Report, TOEFLiBT-05. Princeton, NJ: Educational Testing Service.
- Wall, D. & Horak, T. (2008b). *The impact of changes in the TOEFL examination on teaching and learning in a sample of countries in Europe: Phase 3, The role of the course book. Phase 4, Describing change*. Manuscript submitted for publication, Lancaster University.
- Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal* 13 (2): 7-26.
- Weir, C. (2005). *Language Test Validation: An Evidence-based Approach*. Oxford: Palgrave.

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

Wu, W.M. and Stansfield, C.W. (2001). Towards authenticity of task in test development. *Language Testing* 18 (2): 187-206.

## **APPENDICES**

## Appendix A

### A Needs Analysis Questionnaire

**Dear Sir or Madam,**

Tour guides are one of the key persons in tourism enterprises who deal directly with the linguistic diverse tourists. Therefore, English language literacy is one of the requirements to work as professional inbound tour guides. English speaking skill is essential in their career since they communicate with their customers almost all the time during the organized tour.

Due to the importance of tourism business that creates approximately 6.7% of all the GDP in Thailand in 2010 (Thailand Tourist Arrivals, 2011), a number of educational institutions has offered English for tourism courses which aim to develop tourism professional staff including tour guides who are proficient in the English language. In order to pass the course, students are required to pass the test with certain proficiency. For this reason, a good test is needed to assess the skill precisely and accurately.

As a test developer and researcher, **I would like to receive your opinions regarding the Tourism Knowledge language used tasks, language knowledge and criteria for assessing the language knowledge of the professional English speaking tour guides during tour operation, particularly in the central region from the practitioners' views.** The data collected from this questionnaire will be used as a part of test development in order to derive at the target language use domain. The test tasks are expected to represent the real world tasks and language skills used in the actual context to enable the test developers and users to make valid inference and interpretation on the test scores.

Please fill in the required information on the space provided in this questionnaire. Your opinions and suggestions will be greatly appreciated.

5187805820 Malinee Phaiboonnugulkij

Ph.D. Candidate in English as an International Language Program (EIL),

Graduate School, Chulalongkorn University

Tel: 081-454-6004

E-mail address: [malineerabbit@hotmail.com](mailto:malineerabbit@hotmail.com)

**Needs Analysis Questionnaire for English for Tourism II Test: Tourism Knowledge Language Used Tasks, Language Knowledge and Criteria for Assessing Language knowledge in Tourism Context Focusing on Central Region**

**Part I: Demographic information**

**Instructions: Please fill in the information or put ✓ on the space provided**

**1. Position:** \_\_\_\_\_

**2. Company/Institution:** \_\_\_\_\_

**3. E-mail address:** \_\_\_\_\_

**4. Education**

Diploma Major \_\_\_\_\_

Bachelor's degree Major \_\_\_\_\_

Master's degree Major \_\_\_\_\_

Doctoral degree Major \_\_\_\_\_

Others (please specify) \_\_\_\_\_

**5. Years of experience in tourism profession (please specify) \_\_\_\_\_ years**

**6. Years of experience in English language career**

7-9 years       10-12 years       13-15 years       16-18years

19-21 years       22-24 years       25-27 years       please specify.....

**Part 2: Tourism Knowledge Language Used Tasks**

**Instructions: Please put ✓ in the rating scale for the degree of importance of the following tourism knowledge language used tasks by the tour guides during tour operations and specify additional information on the space provided below.**

4 means very important

3 means important

2 means not so important

1 means not important

Tourism Knowledge Language Used Tasks	Degree of importance			
	4	3	2	1
<b>1. Presenting tourism information</b>				
1.1 Thai history				
1.2 National attractions				
1.3 Thai food				
1.4 Buddhism				
1.5 Architectural structures				
<b>2. Describing tour itinerary</b>				
<b>3. Informing tourists about what they should do and should not do in Thailand</b>				
3.1 At the cultural sites(e.g. the ordination hall)				
3.2 At tour attractions (e.g. shopping center)				
<b>Suggestion for other possible venues:</b> .....				
<b>4. Responding to the following tourists' enquiries</b>				
4.1 Request for help for stolen wallet				
4.2 Want to go out for a night life				
4.3 Ask for help with sick people				
<b>Suggestion: for other possible enquiries:</b> .....				
<b>5. Dealing with tourists' complaints</b>				
5.1 Incomplete tour program e.g. skip one attraction				
5.2 Wrong room request e.g. got smoking room but request for non smoking room				
5.3 Waiting for the bus too long				
<b>Suggestion for other possible complaints:</b> .....				

**Please add other language used tasks that the tour guides performed during tour operations.**

---



### **Part 3: Components of language knowledge for tour guides**

Instructions: Please indicate the degree of importance for the following language knowledge used by the tour guides to conduct the tour.

<b>Importance</b>				<b>Language Knowledge</b>
1	2	3	4	
Not important	Not so important	Important	Very important	
				<b>Knowledge of pronunciation</b> is the ability to use sound, stress and intonation to convey the intended meaning of the utterance.
				<b>Knowledge of vocabulary</b> is the ability to use both generic and tourism related technical terms to respond to the test tasks.
				<b>Knowledge of grammar</b> is the ability to use standard English grammatical structures and rules to produce comprehensible utterances. It includes the use of specific language patterns to construct appropriate responses to the test tasks.
				<b>Knowledge of language functions</b> is the ability to interpret and formulate appropriate and logical speech. It includes the use of heuristic, manipulative, ideational and imaginative functions to respond the test tasks. Knowledge of language function is measured by the appropriateness of the speech produced in terms of the meanings, task requirement and language use setting.
				<b>Knowledge of cohesion</b> is the ability to combine phrases in a meaningful way which can be seen from the use of cohesive devices.
				<b>Fluency</b> is the general quantity of production and tempo of production.
				<b>Content knowledge</b> is the ability to present tourism related content knowledge taught in English for Tourism.

**Is there any additional language knowledge required by the tour guides to communicate with the tourists during tour operations?**

Yes (Please add the language knowledge with a brief explanation on the space below.)

No (Please go to the next part.)

**Part 4: Criteria for assessing language knowledge of the tour guides in speaking ability**

**Instructions: Please put ✓ to give your opinion on the appropriateness of the following criteria to the tour guides' language knowledge in speaking.**

Criteria	Appropriateness			
	4	3	2	1
	Very appropriate	Appropriate	Inappropriate	Very inappropriate
<b>Accuracy</b> is the accurate production of pronunciation, grammatical structure, vocabulary and cohesive devices conforming to standard varieties of English. It includes accurate content knowledge in Thai tourism context.				
<b>Range</b> is the amount of the production in vocabulary, grammatical structures and cohesive devices per utterance.				
<b>Fluency</b> is the amount of utterances produced in a limited time and ability to maintain the pace of rhythm compared to that of native speakers.				
<b>Appropriateness</b> is the appropriate use of grammatical pattern and language function with the consideration of situation, social status and purpose of speech production.				

**Additional criterion/criteria for assessing language knowledge of the tour guides in speaking.**

---



---



---



---



---

Thank you for your cooperation and useful suggestions

\*\*\*\*\*End of the questionnaire\*\*\*\*\*

## Appendix B

### Web-based Speaking Test in English for Tourism (WBST-EFT)

#### Introduction



### Web-Based Speaking Test in English for Tourism (WBST-EFT)

#### Introduction

This is the Web-Based Speaking Test in English for Tourism. The test is the final achievement test that aims to measure your speaking ability in tourism context as required in the English for Tourism II (1552605) course.

In the test, **you will be able to demonstrate your speaking ability as a tour guide conducting the trip in the central region of Thailand. The test comprises of six tasks and lasts approximately 24 minutes.** Respond to each test task as clearly and completely as possible by clicking on the record button and start speaking through the microphone. After you have finished speaking, click off the same button again to stop.

**Please be acknowledged that you are allowed to record your response only once after you have clicked the record button.** Click the next button when you are ready to move to the next section.

In **Task 1**, you will see the pictures and then you will need to explain about the details of each picture by using your tourism content knowledge.

In **Task 2**, you will read the tour itinerary and then you will need to use the given information to talk about the details of the tour program. You will use your content knowledge to explain the attractions.

In **Task 3 and Task 4**, you will listen to the talk by the tour guide from the video clips in different situations. Then, you will see the pop up pictures during the talk. You will need to give the suggestions on the pictures to the tourists. You will use your content knowledge to provide the reasons for your suggestions.


In **Task 5 and Task 6**, you will watch the video clips containing different situations and then you will be faced with both tourists' enquiries and difficulties. You will need to respond to each situation appropriately by expressing your opinion.

**Your response will be judged on your ability to speak clearly, coherently and fluently and your ability to accurately convey the tourism information. You will also be scored on the accuracy, range and appropriateness of vocabulary and grammatical structures used in your response.**

Please listen to the instructions of each task carefully. **You will have a short preparation time in sections 1 and 2.** When the time is up, there will be the **'beep'** sound signaling you to start speaking.

You may take the **sample test** to see what the actual test is like. **Remember this is the sample part so, your response will not be scored!**

 [SAMPLE TEST](#)

 [How to save file?](#)

 [News forum](#)

## Overall page of the WBST-EFT

### 1 Section 1: Presenting Tourism Related Information



(Source: [www.destinationjfo.com](http://www.destinationjfo.com))

 TASK 1

 TASK 2

 Submission Task 1 & Task 2

### 2 Section 2: Giving Polite Suggestions to the Tourists



Do's and  
Don'ts

(Source: [www.bulbrightthai.org](http://www.bulbrightthai.org))

 TASK 3

 TASK 4

 Submission Task 3 & Task 4

3

### Section 3: Dealing with Enquiries and Complaints



(Source: www.koreanfmz.co.kr and www.happytinsipat.blogspot.com)

 TASK 5


 TASK 6

 Submission Task 5& Task 6

4

### The WBST-EFT Attitude Questionnaire



 Attitudes towards the Web-Based Speaking Test in English for Tourism (WBST-EFT) Questionnaire

## Sample test

WBSTEFT: SAMPLE TEST - Windows Internet Explorer

http://lmsonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10746


File Edit View Favorites Tools Help

Search PDFCreator eBay Amazon Coupons Radio Options

Favorites Suggested Sites Web Slice Gallery

WBSTEFT: SAMPLE TEST

Page Safety Tools



Source: www.th.wikipedia.org

Sample test: Listen to the short video clip of the tour guide explaining the attraction. You will see 3 pictures during the talk. Explain the pictures in details. You will have **3 MINUTES** to work on this task. In each picture, you will have **20 seconds** to prepare for your responses and the remaining **40 seconds** to speak.

**Take the actual test!**

Done Internet 100%

Course: We... WBSTEFT: TA... WBSTEFT: TA... WBSTEFT: SA... My Docume... task.doc (3)... EN 14:16


## Task 1: Describing tourist attractions

WBSTFT: TASK 1 - Windows Internet Explorer

http://msonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10743

WBSTFT: TASK 1

### History of the Emerald Buddha Temple



Location  
Year of construction  
Constructor  
Importance of the site  
Important things to see

Task 1; Explain the following pictures in details to the tourists by using your tourism content knowledge. You have **7 MINUTES** to work on the task. In each picture, you may have **20 seconds** to prepare your response and the remaining **40 seconds** will be your speech production.

NEXT

Done Internet 100%

Course: We... Notice: Win... Messages... Document1... WBSTFT: 5... WBSTFT: T...

2:14 PM

## Task 2: Explaining one-day trip tour itinerary

WBSTFEI: TASK 2 - Windows Internet Explorer

http://lmsonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10740

File Edit View Favorites Tools Help

Search PDFCreator eBay Amazon Coupons Radio Options

WBSTFEI: TASK 2

0930 Arrive at Nakhon Pathom

Visit Golden Pagoda and pay respects to the Sacred Buddha Image

1030 Visit Dhavaravati Museum

1130 Sightseeing and buy souvenirs at the Local Market

1245 Have lunch at 'Ban Ruen Thai Restaurant', a famous restaurant in Nakhon Pathom that offers delicious local dishes such as grilled river prawns, spicy salad with fresh squid, fried chicken with Thai herbs, and spicy Nakhon Pathom soup

1345 Visit Sanam Chandra Palace

1500 Depart from Nakhon Pathom

1630 Arrive safely at The Grand Hotel, Bangkok

(Source: www.Tourthai.ky.com)

**Task 2:** Read the following one-day trip itinerary. Use the given information to explain this tour program. You are required to add the details concerning the underlined attractions. You have **5 MINUTES** to work on the task which is **2 minutes** for reading and organizing your idea. The remaining **3 minutes** will be for your speech production. The details of the itinerary are as follows.

**NEXT**

Done Internet 100% 13:59

Course: Web-Based... WBSTFEI: TASK 4 - ... Document1 - Micros... WBSTFEI: TASK 2 - ...



### Task 3: Giving polite suggestions to the tourists at the ritual sites

WBSTFT: TASK 3 - Windows Internet Explorer

http://msonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?id=10737

WBSTFT: TASK 3



Source: www.highsnobiety.com, www.bang.postcard.com, www.inz.flickr.com

Task 3: Listen to the tour guide's talk from the video clip. There will be 6 pop up pictures in order. Give appropriate suggestions with reasons to each picture. You will have 3 MINUTES to work on this task. For each picture there will be 10 seconds for organizing your idea and the remaining 20 seconds to speak.

NEXT

Done Internet 100%

Course: Web-Based S... Notice - Windows Inte... Document1 - Microsoft... WBSTFT: TASK 3 - W... 2:15 PM

## Task 4: Giving polite suggestions to the tourists at the crowded attractions

WBSTFEIT: TASK 4 - Windows Internet Explorer

http://msonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10738


File Edit View Favorites Tools Help

pdfforge Yahoo Search PDFCreator eBay Amazon Coupons Radio Options

Favorites Suggested Sites Web Slice Gallery

WBSTFEIT: TASK 4

Page Safety Tools



(Source: www.accountrements.com)

Task 4: Listen to the tour guide's talk from the video clip. There will be 6 pop up pictures in order. Give appropriate suggestions with reasons to each picture. You will have 3 MINUTES to work on this task. For each picture there will be 10 seconds for organizing your idea and the rest 20 seconds to speak.

NEXT

Internet 100%

Course: Web-Based... WBSTFEIT: TASK 4 - ... Document1 - Micros... EN 13:59

## Task 5: Dealing with tourists' enquiries

WBSTEFT: TASK 5 - Windows Internet Explorer

http://msonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10751

WBSTEFT: TASK 5



Task 5: Watch 3 short video clips on different enquiries. Give appropriate responses to each situation. You need to speak right after the signal. **THERE WILL BE NO PREPARATION TIME ON THIS TASK.** You will have 3 minutes to speak that is 1 minute for each situation.

**NEXT**

Done Internet 100%

Course: Web-Based S... Notice - Windows Inte... Document1 - Microsoft... WBSTEFT: TASK 5 - W... 2:17 PM

## Task 6: Dealing with tourists' complaints

WBSTEF: TASK 6 - Windows Internet Explorer

http://msonline2.nrru.ac.th/NRRU-Online/mod/resource/view.php?inpopup=true&id=10752

File Edit View Favorites Tools Help

Search PDFCreator eBay Amazon Coupons Radio Options

Favorites Suggested Sites Web Slice Gallery

WBSTEF: TASK 6

Page Safety Tools



**Task 6:** Listen to the 3 tourists' complaints from the video clips. Give appropriate responses to each complaint. You need to speak right after the signal. **THERE WILL BE NO PREPARATION TIME ON THIS TASK.** You will have 3 minutes to speak that is 1 minute for each situation.

Done Internet 100%

Course: Web-Based... WBSTEF: TASK 4 - ... Document1 - Micros... WBSTEF: TASK 6 - ... EN 14:00

## Appendix C

### The Rating Scale

#### Construct definitions

The construct definitions were derived from the frameworks of Douglas (2000) for some language knowledge and Fulcher (2003) for speaking ability. All the definitions were operationally based on the course syllabus of English for Tourism II.

**Knowledge of pronunciation** is the ability to use sound, stress and intonation to convey the intended meaning of an utterance. It is measured by the degree of accuracy to pronounce words that conform to standard varieties of English. It also includes the effective use and degree of intelligibility of stress to emphasize particular words and use of intonation to convey speech functions. The average score from two raters is the representation of this knowledge.

**Knowledge of vocabulary** is the ability to use both generic and tourism-related technical terms to respond to the test tasks. It is measured by the accuracy and range of the vocabulary employed in the responses. The knowledge of the vocabulary is demonstrated through the average score from two raters.

**Knowledge of grammar** is the ability to use standard English grammatical structures and rules to produce comprehensible performances. It includes the use of specific language patterns to construct appropriate responses to the test tasks. It is assessed by the accuracy, range, complexity and appropriateness of the structures in the speech produced. The average score from two raters is the representation of this knowledge.

**Knowledge of language functions** means the ability to interpret and formulate appropriate and logical speech. It includes the use of ideational, manipulative, heuristic and imaginative functions to respond to the test tasks. Knowledge of language functions is measured by the appropriateness of the speech produced in terms of the meanings, task requirements, and language use setting. The average score from two raters is the representation of this knowledge.

**Knowledge of cohesion** is the ability to combine phrases and sentences in a meaningful way, which can be seen from the use of cohesive devices in the responses. It is measured by the accuracy and range of cohesive markers in the test performances. Knowledge of cohesion is demonstrated through the average score from two raters.

**Fluency** is the general quantity and tempo of language production. It is the ability to use the tempo and pauses in language production to maintain paces of the responses. It is assessed by the appropriate use of both the tempo and pauses in the responses. Fluency of speech is demonstrated through the average score from two raters.

**Content knowledge** is the ability to present tourism-related content knowledge taught in English for Tourism II. It is measured by the accuracy and completion of the information given by the students to respond to the test tasks. Content knowledge is demonstrated through the average score from two raters.

**WBST-EFT Rating Scale**

	<b>Bands</b>	<b>Descriptors</b>
<b>Pronunciation</b>	4 Very good	Occasionally mispronounce words but do not affect the intelligibility of the utterance. Effective and appropriate use of stress and intonation to emphasize the meaning and function of speech
	3 Good	Frequently mispronounce words but do not interfere the meaning of the speech. Appropriate use of stress with and intonation to emphasize the intended meaning and achieve particular purposes
	2 Fair	Constantly mispronounce words that make the utterance hard to understand. Occasional use of stress and intonation to emphasize particular meaning and convey limited speech functions
	1 Beginner	Almost mispronounce words which cause confusion of the utterance. Inappropriate use of stress and intonation that interfere with the intended meaning and speech functions
	0 Very poor	Pronounce words mostly in an incomprehensible way. Speech is very hard to understand and no evidence of using stress and intonation

	<b>Bands</b>	<b>Descriptors</b>
<b>Vocabulary</b>	4 Very good	Wide range of vocabulary with precise meaning of technical terms. Evidence of very few circumlocutions in certain complex field specific topics. Few errors but do not affect the intelligibility of the speech
	3 Good	Adequate range of technical terms to convey the meaning. Evidence of some circumlocutions in certain complex field specific topics. Few errors but do not hinder the intelligibility of the speech
	2 Fair	Uses of technical terms are sometimes limited and inaccurate that hinder the overall comprehension of the speech in some topics
	1 Beginner	Limited range of technical terms. Frequent use of inaccurate word choice and forms that cause misunderstanding and confusion in the utterance
	0 Very poor	Inadequate and inaccurate use of technical terms to convey even simple construction



	<b>Bands</b>	<b>Descriptors</b>
<b>Grammar</b>	4 Very good	Use of complex and variety of constructions with few errors that do not cause any confusion and unintelligibility of the speech. Appropriate use of grammatical patterns on specific social situation that reflect the control of the major rules
	3 Good	Complex constructions with occasional errors that represent most control of major grammatical structures. No effect of errors on the intended meaning of the speech. Appropriate use of certain pattern for particular social situation
	2 Fair	Simple construction with frequent errors that represent some knowledge of grammatical rules. No evidence of errors causing major confusion and misunderstanding of the utterance
	1 Beginner	Short and very simple construction with constant errors that obstruct the comprehension of the speech, accurate use only stock phrases
	0 Very poor	Evidence of almost ungrammatical structures particularly on the word orders that are unable to understand. Accurate use of the rules only in formulaic expressions

	<b>Bands</b>	<b>Descriptors</b>
<b>Language functions</b>	4 Very good	Appropriate use of language functions in regards to meaning and language use setting. Effectively respond to task requirements with minimal irrelevant information that does not affect the appropriateness of speech produced
	3 Good	Appropriate use of language functions on the aspect of meaning and language use setting. Respond to task requirements with minor insufficient and irrelevant information
	2 Fair	Minimal use of appropriate language functions in terms of meaning and language use setting. Respond partially to many of task requirements with major problems in giving irrelevant and inadequate information
	1 Beginner	Only able to respond partially to a few of task requirements
	0 Very poor	No evidence of appropriate language functions

	<b>Bands</b>	<b>Descriptors</b>
<b>Cohesion</b>	4 Very good	Constant use of cohesive markers to combine and contrast ideas in phrases and sentences to make a meaningful speech. Minor errors but do not affect the overall comprehensibility and coherence of the utterance
	3 Good	Frequent use of cohesive devices to connect or contrast ideas in phrases and sentences to make the speech comprehensible. Few errors but do not interfere with the meaning of the speech
	2 Fair	Occasional use of cohesive markers in phrases and sentences with major errors in the meaning of the overall idea
	1 Beginner	Very few and mostly incorrect use of cohesive devices in phrases and sentences
	0 Very poor	No evidence of using cohesive devices in phrases and sentences

	<b>Bands</b>	<b>Descriptors</b>
<b>Fluency</b>	4 Very good	Almost effortless with few obstructions from hesitation. Very few lexical repetitions and rephrasing
	3 Good	Occasional hesitation with fillers and non-fillers but do not obstruct the understanding of the speech. Some uneven speeches from word grouping and restructuring of the phrases
	2 Fair	Uneven speeches with frequent pauses that obstruct the overall intelligibility. Lexical item repetitions and rephrases with incomplete sentences
	1 Beginner	Uneven and disconnected speech with frequent pauses except stock phrases
	0 Very poor	Disconnected and uneven speech that is unable to understand

	<b>Bands</b>	<b>Descriptors</b>
<b>Content Knowledge</b>	4 Very good	Accurate and precise information. Evidence of very few incomplete information except in very complex topics
	3 Good	Accurate and precise information with minor errors. Evidence of some incomplete information particularly in the complex topics
	2 Fair	Few errors with fairly precise information. Some missing information as required from the test tasks
	1 Beginner	Some errors with broad information. Evidence of very few related but with much incomplete information in speech
	0 Very poor	Inaccurate, too general and unrelated information as required from the test tasks

## Appendix D

### The Attitudes towards the WBST-EFT Online Questionnaire



## Attitudes towards the Web- based Speaking Test in English for Tourism (WBST-EFT) Questionnaire

The aim of this questionnaire is to investigate your attitudes towards the Web-Based Speaking Test in English for Tourism (WBST-EFT). Your information will help us measure the overall usefulness of the test in assessing the speaking ability in Tourism oriented context.

Part of your opinions will be used to improve this innovative test. Your opinions and suggestions will be valuable information for us to develop this protocol test. Please fill in the required information on this questionnaire with the detailed explanation. Thank you for your cooperation.

\*จำเป็น

**Part I: Demographic Information**

Instructions: Please fill in or check the answer

**1. Gender \***

- Male
- Female

**2. Faculty and Major \*****3. Group \***

- A1
- A2
- A3
- B1
- B2
- B3

**4. Have you ever taken any internet-based or computer-based test? \***

- Yes
- No

**Part II: Opinions on the WBST-EFT**

Instructions: Please rate your opinion on the WBST-EFT for the following statements.

1 means Strongly disagree

2 means Disagree

3 means Agree

4 means Strongly agree

**1. Overall usefulness**

**1.1 I feel the WBST-EFT provided me the adequate opportunity to demonstrate both of my strengths and weaknesses on speaking ability.**

1 2 3 4

Strongly disagree     Strongly agree

**1.2 I think the raters listening to my response via the WBST-EFT will get an accurate idea of my speaking ability in tourism context as stated in the course syllabus.**

1 2 3 4

Strongly disagree     Strongly agree



**1.3 I felt at ease taking the WBST-EFT.**

1 2 3 4  
Strongly disagree     Strongly agree

**1.4 The instructions are clear and easy to follow.**

1 2 3 4  
Strongly disagree     Strongly agree

**1.5 The introduction part is useful because it gives me the example and chance to practice the test.**

1 2 3 4  
Strongly disagree     Strongly agree

**1.6 The tasks and situations on the WBST-EFT are appropriate and simulate the real world tasks**

1 2 3 4  
Strongly disagree     Strongly agree

**1.7 The test taking procedures are too sophisticated for me and require proficiency in computer.**

1 2 3 4  
Strongly disagree     Strongly agree

**1.8 The Web-based test is not an appropriate test for speaking ability.**

1 2 3 4  
Strongly disagree     Strongly agree

**2. Appropriateness of time for preparation and response formulation**

**2.1 The preparation time for Section1(Presenting tourism related information) is adequate.**

1   2   3   4

---

Strongly disagree               Strongly agree

**2.2 The preparation time for Section 2(Asking tourists with polite requests) is adequate.**

1   2   3   4

---

Strongly disagree               Strongly agree

**2.3 The time allowed for response formulation for Section1 is appropriate.**

1   2   3   4

---

Strongly disagree               Strongly agree

**2.4 The time allowed for response formulation for Section2 is appropriate.**

1   2   3   4

---

Strongly disagree               Strongly agree

**2.5 The time allowed for response formulation for Section3 is appropriate.**

1   2   3   4

---

Strongly disagree               Strongly agree

### 3. Task difficulty

**3.1 I think the tourism related information tasks are too difficult.**

1   2   3   4

---

Strongly disagree     Strongly agree

**3.2 I think asking tourists with polite requests is too difficult.**

1   2   3   4

---

Strongly disagree     Strongly agree

**3.3 I think dealing with enquiries and complaints are too difficult.**

1   2   3   4

---

Strongly disagree     Strongly agree

### 4. Interface design

**4.1 I think the navigation button, icon (e.g. recording and move buttons), tool bar, and controls are easy to use.**

1   2   3   4

---

Strongly disagree     Strongly agree

**4.2 The terminology used in the test is hard to understand.**

1   2   3   4

---

Strongly disagree     Strongly agree

**4.3 Each page layout contains appropriate detail, clear title and is easy to read.**

1   2   3   4

---

Strongly disagree     Strongly agree

**4.4 I think the text, font size, and color used in the test are appropriately designed.**

1 2 3 4  
Strongly disagree     Strongly agree

**4.5 The multimedia (e.g. video clips & audio files) used in this test are appropriate.**

1 2 3 4  
Strongly disagree     Strongly agree

**4.6 The multimedia help me understand the prompt better and do not take too long to download.**

1 2 3 4  
Strongly disagree     Strongly agree

**4.7 I could use the help facilities (e.g. pop up explanation) while I was taking the test.**

1 2 3 4  
Strongly disagree     Strongly agree

### Part III: Open-ended questions

Instructions: Please answer the following questions with detailed explanations on the space provided below.

**1. What do you consider the strengths and weaknesses of the WBST-EFT?**

**2. In your opinion, please identify what you liked and did not like most about the test. Please give the reasons.**

**Thank you for your cooperation and useful suggestions**

\*\*\*\*\*End of the questionnaire\*\*\*\*\*

ส่ง

สนับสนุนโดย [Google Documents](#)

รายงานการละเมิด - ข้อกำหนดในการให้บริการ - ข้อกำหนดเพิ่มเติม

## Appendix E

### List of Speaking Test Taking Strategies, Speaking Test Taking Strategies Elicitation Instructions, and Interview Scripts and Questions

#### List of Speaking Test Taking Strategies

The taxonomy of the speaking test taking strategies was derived from Bachman & Palmer (1996), Fulcher (2003), Cohen (1998) and Swain et al. (2009).

**1. Communication strategies** deal with conscious planning to solve linguistic difficulty during communication. Strategies in this category are as follows:

**1.1 Achievement strategies** are used when the test takers face the communicative problem due to the lack of language knowledge.

- Overgeneralization/morphological creativity: transferring the existing linguistic rules to the unknown rules when having linguistic difficulty
- Approximation: replacing more generic terms to the unknown words
- Paraphrase: using circumlocution or near synonym words
- Word coinage: creating new words for the unknown words
- Restructuring: using different grammatical structures and words to convey the same message when they think the sentence is unable to be understood
- Code switching: using words or phrases from the native language with the target language when they lack target language linguistic knowledge

**1.2 Avoidance strategies** are used when the learners lack control over the target language.

- Topic avoidance: avoiding the unknown topic
- Conversation abandoning: leaving the message incomplete
- Semantic avoidance: avoiding the unknown with the use of generic words e.g. things

**2. Cognitive strategies** deal with the target language manipulation in understanding and producing the language. Four strategies are included as follows:

**2.1 Selecting or attending strategy** is used to direct the attention to a specific feature of the task.

**2.2 Comprehending strategy** deals with seeking ways to understand the task. This strategy consists of various strategies which are clarifying information, L1 translating, inference, analyzing, and reasoning.

**2.3 Storing memory strategy** deals with finding ways to memorize information to respond the test tasks. It includes repeating on what is read or heard, associating the existing information, linking new information with previous knowledge, summarizing the L2 information and using imagery to generate, understand or remember the information.

**2.4 Retrieval strategy** is used to retrieve L2 linguistic resource and background knowledge to attempt the test tasks. It includes recombining the target language knowledge, transferring L1 linguistic knowledge, translating L1 to L2 and recalling.

**3. Metacognitive strategies** deal with the conscious investigation of the test taking process to identify, organize, evaluate and plan the effective ways in attempting the test tasks. It is made up of three sub-strategies.

**3.1 Goal setting strategy:** identifying the tasks and deciding what to do

**3.2 Assessment strategy:** assessing what is needed, what one has to work with and how well one has done

**3.3 Planning strategy:** deciding how to use the existing language knowledge and background knowledge

**A. Instructions for collecting test taking strategies used in attempting the test tasks**

1. Make the participants feel at ease.
2. Inform the participants about the objective of this session.
3. Give clear instructions on what the participants are required to do.
4. Check whether the web camera is working.
5. Ensure that the response is from the participant not from your guidance. Do not go further than asking “what were you thinking?”.
6. In case that the participants cannot remember, do not pursue to get the response.

7. Interact flexibly with the participant by using back channeling e.g. uh huh, I see, Okay. Do not give definite reaction to the response.
8. Make sure to pause the video clip while the participant is giving the response during the recall session.

### **B. Interview script and questions**

The objective of this session is to know what were you thinking while performing the semi-direct six test tasks that were administered via the internet system. Please be informed that your response will be recorded.

When you complete each test task, I am going to play back the video clip and ask you to report your thought before, during and after the task performance. This also includes whatever that comes to your mind during the playback. The important thing is please keep talking and do not plan to explain what you are thinking. You can keep talking in English or in Thai. If you stop talking or remain silent for any moment, I will remind you to keep talking.

O.K. Do you have any questions about what I have asked you to do?

Next, we are going to watch the video clip that you have performed the task. Please remember to report anything in your mind before, during and after you completed each test task. Please keep talking but do not report on what you may have or should have done during the task performance. Please also keep in mind that I do not want you to do the test again. While you are watching the video clip, you can pause anytime to report your thought. And if I have any questions, I will pause the clip and ask you to clarify the point for me. Please start and tell me what you can remember.

Sample general questions to ask during the strategies report session.

- At this point, what were you thinking?
- Can you tell me what you were thinking when you.....?
- Is there anything else that you can remember?
- What came up in your mind at this point?
- You look a bit confused. What were you thinking of?
- You look..... What were you thinking then?



In case that the subjects do not speak at all, use more specific questions to elicit the strategies used. The questions are as follows:

**Achievement strategies**

- What did you do when you could not continue your speech?
- Did you apply a particular structure to others?
- Did you use general term with the unknown word?
- Did you use the word that had close meaning to the one you could not think of or create a new one?
- Did you rearrange your sentence to convey the same meaning as the previous one?
- Did you switch to your native language?

**Avoidance strategies**

- When you could not continue your speech did you try to avoid the unknown topic or just leave the message incomplete?
- Did you skip the unknown word and use the one that had similar meaning?

**Selecting strategy**

- How did you formulate your response to the test tasks?
- Did you focus on a particular feature of the task?

**Comprehending strategy**

- How did you understand the task?
- Did you clarify the test prompt?
- Did you translate the test prompt to your native language?
- Did you analyze the test prompt or try to find the reason behind it?

**Storing memory strategy**

- How did you memorize information on the test tasks?
- Did you repeat on what you have read or heard?
- Did you relate the existing information to the new one?
- Did you link new information with previous knowledge?
- Did you summarize information on the test tasks?
- Did you use any imagery to generate, understand or remember the information?

**Retrieval strategy**

- When you want to formulate your response did you retrieve L2 linguistic knowledge and your subject specific knowledge?
- Did you recombine your L2 language knowledge or transfer your L1 linguistic knowledge?
- Did you translate your response into your native language?
- Did you recall your response?

**Goal setting strategy**

- What were your thinking processes when you attempted the test tasks?
- Did you identify the tasks and decide what to do?

**Assessment strategy**

- Did you assess what was needed, what did you have to work with and how well you have done?

**Planning strategy**

- Did you decide how to use the existing language knowledge and background knowledge?

## Appendix F

### Samples of the Speech Performances and the Test Scores

#### A1.1 (Task type one: High proficiency group)

##### Task 1:

Picture1: This is the Emerald Buddha Temple. It was built in the reign of King Rama the first in 1782. The Emerald Buddha Temple was very important because Thai people believed that it was the most sacred place[s] in Thailand and the repository of spirits for all Thai people. There are many interesting things to see inside for example the Emerald Buddha image, the ordination hall and the gallery.

Picture 2: Now the emerald Buddha image is enshrining inside the ordination hall at the Emerald Buddha Temple. This is the Emerald Buddha image. It was made from green jade and it is 48.3 cms wide and high. The Emerald Buddha image is [in] meditation posture. It has 2 seasonal costumes made in the reign of King Rama the 1<sup>st</sup>, one for summer and one for the rainy season. Then, in the reign of King Rama the 3<sup>rd</sup> in 1824 he add[ed] another one for winter.

Picture3: We are now standing in front of the ordination hall. It was built by King Rama the 1<sup>st</sup> in 1782. The ordination hall is very important because it houses the emerald Buddha image which is the most sacred image in Thailand. There are many interesting things to see inside for example the emerald Buddha image, พระพุทธสัมปณันี image, the murals and outside you can see the base of the ordination hall are garudas holding nagas.

Picture4: Gallery. The gallery that you can see was built by King Rama the 1<sup>st</sup> in 18<sup>th</sup> centuries. Frequent restorations were carrying out during their century of existence mainly during the 3<sup>rd</sup>, 4<sup>th</sup> and the 4<sup>th</sup> reign. The most recent restoration was started in 1975 and it took 7 years to complete. The walls are decorate[ed] with murals that describe the entire Ramakien or Ramayana which have 178 scenes.

Picture5: This is the Grand Palace. It was built in 1782 by King Rama the 1<sup>st</sup>. [The] original living quarters were temporally made of wood and thatch. After coronation, King Rama the 1<sup>st</sup> built of permanent building. There are many interesting things to see inside such as Phra Maha Monthain Group, Phra Maha Prasat Group, Phra Thinang Chakri Maha Prasat Group, Borophiman Mansion and Siwalai Garden Group

Picture 6: This is the Phra Thinang Chakri Maha Prasat Group [Chakri Maha Prasat Thronehall]. This building was built by King Rama the 5<sup>th</sup> or King Chulalongkorn to commemorate the centenary of the Chakri dynasty. It was designed by a British architect in European style and pure Thai style roof. Construction took 6 years from

1876 to 1882. There are many interesting things to see inside such as the galleries, the libraries and the Chakri Throne Room.

Picture7: This is Dusit Mahaprasat Throne Hall. King Rama the 1<sup>st</sup> built Dusit Mahaprasat as a replacement for an earlier wooden พระที่นั่งอัมรินทร์ภิเษกมหาปราสาท which burn [t] down in 1790. Dusit Mahaprasat Throne Hall is very important because it [was] used for the annual consecration ceremony or พระราชพิธีถือน้ำพระพิพัฒน์เกล้า. There are many interesting things to see inside such as พระราชบัลลังก์ประดับมุก.

## Task 2:

Bangkok-Nakhon Pathom

First, at 08 o'clock we will depart from the Grand Hotel Bangkok.

Second, at 0930 we will arrive at Nakhon Pathom and visit the Golden Pagoda and pay respect to the scared Buddha image.

Third, at 1030 we will visit ทวารวดี Museum.

Next, at 1130 we will sightseeing and buy a souvenir[s] at the Local market.

After that have lunch at 'Ban Ruen Thai Restaurant', a famous restaurants in Nakhon Pathom, that offers the delicious local dish[es] such as [pause]grilled[gill] river pround [prawn], spicy salad[pause] with fred [fresh] sguir[squid], fried chicken with Thai herbs, and spicy Nakhon Pathom soup at 1245.

After that we will visit 'Sanam[pause]chandra Palace'[place] at 1345.

And the lart [last] we will depart from Nakhon Pathom at [pause] fifty p.m. [fifteen o'clock].

Finally, we will arrive safely at The Grand Hotel Bangkok at [pause] forty thirty fourteen thirty.

Task	Pronun	Voc	Gram	Lang Func	Cohesion	Fluency	Cont. know
1	3	4	4	4	4	4	4
2	3	4	4	4	4	4	3

### B1.1 (Task type one: Low proficiency group)

#### Task 1:

Picture 1: This is the Emerald Buddha Temple. It was built in the reign of King [King] Rama [the] first. It was built by King Rama the first. It is the symbol of Thai nation. There are many interesting [things] to see for example the ordination [pause] hall and the Emerald Buddha image [image].

Picture 2: Now the Emerald Buddha image [image]. It was made from jade. It is in the posture [posture] of meditation. It is 48.3 cms wide and 66 cms tall. It was found in เชียงราย in the north [n] [of] Thailand. It is enshrined in the ordination hall.

Picture 3: This is the ordination hall. It was built in the reign of King [King] Rama [the] first. It was built by King Rama the first. It contains the Emerald Buddha image [image]. There are many interesting [things] to see for example the best [base] of the ordination hall and garudas holding naga. Several murals [murals] and the gallery outside.

Picture 4: This is the gallery. It was built in the reign of King [King] Rama [the] first. It was built by King Rama the first. It is noted for its murals depicting the entire รามเกียรติ์. There are many interesting [things] to see for example the inside wall decorated with mural [murals] that depict[s] entire Ramayana and the 178 scenes believe at note great and continuing clockwise.

Picture 5: This is the Grand Palace. It was built in 1782. It was built by King Rama the 1<sup>st</sup>. It is a residence [residence] of King Mongkut Rama the third until the entire age of his life. There are many interesting [things] to see for example the พระมหามณเฑียร group, พระมหาปราสาท group, พระที่นั่งจักรีมหาปราสาท group, กรมพิมาน Mansion and ศาลาย group.

Picture 6: This is the พระที่นั่งจักรีมหาปราสาท. It was built in the reign [reign] of King จุฬาลงกรณ์. It was built by King Rama the fifth. It is a reception hall for royal great [guest]. There are many interesting [things] to see for example in the rear center of จักรีมหาปราสาท is the Chakri Throne [Throne] room and the symbol of Chakri dynasty [dynasty] is present [present] on the wall behind the throne [throne].

Picture 7: This is the Dusit Mahaprasat Throne [Throne] Hall. It was built in seven [pause] 17 [pause] 89. It was built by King Rama the 1<sup>st</sup>. It used for annual



**A2.1 (Task type two: High proficiency group)****Task 3:**

Please don't put shoes inside because it is impolite in Thai culture.

Please don't wear spaghetti shirt [singlet] and jean short[s] inside because it is impolite.

Please don't take photo in the ordination[ordination] hall because it is the regulation.

Please don't make noit [noice] when you inside the ordination[ordination] hall because it will bother other people.

Please don't ciming[climbing] on the Buddha image because it will endtrust[distrust] Thai people belief if you offend[ed] in Buddhism.

Please don't sit on the floor of the temple with your feet pointing at the Buddha imesh[image] because it is consider[ed] as highly impolite in Thai culture.

**Task 4:**

Please you should beware as[of the] pickpocket because there are people in the area.

Please call one one five five if you have any pro\_blem [problem] because [pause] the police can help you.

Please go to follow the map because it will lead you [to] the hotel.

Please don't bring belong\_ging[belonging] inside because it is the regulation.

Please be punctual because maybe you miss trip.

Please [use] free service to [at] the JJ Mall near the tourist police office because tourist police[s] in Thailand want [to] take care everybody.

Task	Pronun	Voc	Gram	Lang Func	Cohesion	Fluency	Cont. know
3	3	3	3	4	3	3	3
4	3	3	3	4	3	3	3

**B2.6 (Task type two: Low proficiency group)****Task 3:**

Please you take off your shoes [pause] becaud [because] becaud [because] not it is to it is ruse[rule] มันเป็นกฎ.

Please do not[wear] spaghetti shirt[singlet] [pause] [because] it is the ruse[rule].

Please do not take please do not take a photo becaud [because] it is the ruse[rule].

Do not stop mouth [pause] becaud [because] polite becaud [because] it is polite.

Please do not walk on the yars [yard][pause] becaud [because it is the ruse[rule].

Please sit down Please sit down.

**Task 4:**

Please beware [pause].

Tourist police call becaud [because][pause].

Please cat[carry] map becaud [because you will] not get lot[ lost].

Please do not belonging becaud [because].

Please do not times แล้วยก becaud [because] not miss the bud [bus].

Tour tour Jatujak.

Task	Pronun	Voc	Gram	Lang Func	Cohesion	Fluency	Cont. know
3	1	1	1	2	1	1	1
4	1	1	1	2	1	1	0



### A3.1 (Task type three: High proficiency group)

#### Task 5:

Sir I will provide the tour program that you want right now. Our tourist program is always available for you all the time. Proving your trip, facility and other to you. You can we can provide you [pause] to go [pause] somewhere in Bangkok [pause] for example uh temple [pause] the Emerald Buddha.

I'm sorry. I will give the medicine for your son right now. Your son shouldn't drink cool water. I will try my best.

Oh I'm sorry to hear that. I will take you to the [pause] tourist police right now. Maybe your wallet [pause] lie at some place.

#### Task 6:

I'm sorry I will contact my bus right now. I tried to contact the driver be late be on time next time and I won't let it happen again.

I'm sorry madam. I will contact housekeeping right now. The housekeeping will take care of you. And I won't let it happen again next time. And everything you [pause] everything that happen is ok.

Certainly, that's no problem if you want we will go visit and shopping at the floating market. You will see various handicrafts and can buy some souvenir and buy some food in Ayutthaya or you can sightseeing in the temple and statue in the temple and beautiful scenery in Ayutthaya at Ayutthaya. Maybe information is not enough for today I apologize for you. I will find information for you next time.

Task	Pronun	Voc	Gram	Lang Func	Cohesion	Fluency	Cont. know
5	4	4	4	4	4	4	4
6	4	4	3	4	4	4	4

**B3.18 (Task type three: Low proficiency group)****Task 5:**

OK will\_ tow[tell] you\_ now. Jack[Just]\_ moment\_ please. I [pause]call[pause] to[pause] the driver now.

Uh\_ OK \_will\_ tow[tell] you now. Jack[Just] moment please. I[pause] call[pause] to[pause] the driver now.

Uh\_OK\_ I\_ will\_ call\_ the\_ doctor\_ now. Please [pause] give\_ me\_ the\_ doctors.

**Task 6:**

OK\_ I'm\_ sorry. The\_ place\_ now\_ for\_ nice\_ please.

OK\_ I'm\_ sorry [pause]. Excuse\_ me [pause] please.

Uh\_ OK\_ I'm\_ sorry. I [pause] will[pause] help\_ you\_ uh.

Uh\_ I'm\_ sorry [pause] I [pause] will[pause] help\_ the\_ driver\_ now[pause] please.

Task	Pronun	Voc	Gram	Lang Func	Cohesion	Fluency	Cont. know
5	0	1	1	0	0	0	0
6	0	0	0	0	0	0	0

## Appendix G

### The Speaking Test Taking Strategies Coding Scheme

Strategy/sub-strategies	Definitions	Examples
<b>1. Communication strategies deal with conscious planning to solve linguistic difficulty during communication</b>		
<b>1.1 Achievement strategies</b> are used when the students face the communicative problem due to the lack of language knowledge.		
<b>Approximation</b>	Students using more general term for the unknown word	<p>L1.1: [I] talked in general like I used “Place” instead of “Pavilion” because I could not remember that vocabulary. I could not think of it.</p> <p>L2.1: [I] used of the vocabulary like “clothes” for shorts. I could not remember that word.</p>
<b>Paraphrase</b>	Students using circumlocution or near synonym words	<p>H1.1: I substituted the difficult vocabulary and used the word that I knew and had the closest meaning.</p> <p>H2.3: [I] used the vocabulary that I learned in the class for the unknown words in the picture. For example “a place to see the Buddha image” was used for “พระอุโบสถ [the ordination hall]”.</p> <p>H3.3: [I] tried to find the near synonym word like “Bus to BTS” for “รถรับส่งผู้โดยสาร [Transfer bus]”.</p> <p>L1.3: [I] could not think of any words. [I] used the one that could think of. [I] used “Sitting Buddha image” for “พระพุทธรูปปางมารวิชัย [Subduing Mara Buddha image]”.</p>

		L2.2: Sometimes [I] used the near synonym word. For example [I] used “Bad sitting position” to substitute “นั่งไม่สุภาพ [Sitting in impolite position]”.
		L3.2: [I] used easy and known vocabulary that [I] could remember at that time. The example was [when the tourist wanted to] change the room [I] used “Hotel person at the counter” for “พนักงานจองโรงแรม [Reservationist]”.
<b>Word coinage</b>	Students creating new words for the unknown words	H3.1: [I] created the new word. [I] did not know the word “คนขับรถทัวร์ [Bus driver]” because there was no preparation time. [I] needed to answer right the way so [I] used “Chauffeur driver”.
<b>Restructuring</b>	Students using different grammatical structures and words to convey the same message when they think the sentence is unable to be understood	H3.2: [I] used easy sentences that [I] could think of to answer [the question].  H3.3: [I] used the closest meaning sentence like “Sorry sir I hope it’s not happens again” for “I’m sorry sir I won’t let it happens again”.
<b>Code switching</b>	Students using words or phrases from the native language with the	H1.1: [I] answered in Thai because I could not come up with any English words. [I] used “the พระที่นั่ง”[the Throne Hall] to answer.

---

target language  
when they lack  
target language  
linguistic  
knowledge

---

**1.2 Avoidance strategies** are used when the students lack control over the target language

---

<b>Topic avoidance</b>	Students avoiding the unknown topic	<p>H2.2: [I] skipped the topic and continued with the new picture.</p> <p>L1.2: If [I] cannot remember the information [I] skip it for example, who was the constructors?</p> <p>L2.2: [I] skipped the details on [the reason] why it was prohibited to wear sleepless shirt inside the temple.</p>
<b>Conversation abandoning</b>	Students leaving the message incomplete	<p>H1.3: [I] gave short answer and skipped to the next item. It was better than did not say anything at all. “This is Chrackri Mahaprasart and...”.</p> <p>L1.3: [I] skipped to the next question but I did say something. พระอุโบสถ [the ordination hall] is...</p> <p>L2.2: [I] skipped to the next one if [I] could not answer for; example, “Please don’t take off your shoes because...”.</p> <p>L3.2: [I] skipped. For example, the reason that the tour program was changed. [I] could not come up with any reasons. I could not think of anything...</p>

---

---

## 2 .Cognitive strategies deal with the target language manipulation in understanding and producing the language

---

**2.1 Selecting or attending strategy** is used to direct the attention to a specific feature of the task

---

Students focusing on the salient features of the pictures	<p>H1.1: [I] looked at the picture, time and places and set my own answer. [I] arranged the information [that which one should] come before or after to answer the question based on the existing information.</p> <p>H2.3: [I] mainly looked at the pictures and thought of the vocabulary from the class. For example, the picture of do not wear [not wearing] shoes [I] recognized the salient feature of each picture. Picture with prohibition usually has the cross.</p> <p>L1.1: [I] looked at the picture [and figured out] where the place was.</p> <p>L3.1: [I] focused on the situation, pictures and captured on tourists' speech...</p>
---	---

---

Students focusing on important information and key words	<p>H1.3: [I] focused on the important part in my own words. [To explain] วัดพระแก้ว [the Emerald Buddha Temple], [I] used “This is the Emerald Buddha Temple.”</p> <p>L1.1: ... [I] looked for the key words. [I] tried to find the focus of the place [so that I could] speak a lot. For another task [I] emphasized on time and place.</p> <p>H2.2: [I] focused on the importance of the task... For example, if [I] want to warn the tourist for not doing something [I] used “You should not climbing [climb] the Buddha image”.</p>
--	--

---

	Students emphasizing on the instruction	H3.1: [I] read the instruction of the task and tried to capture on what the tourists wanted.
	Students emphasizing on the types of questions from tourists' speech	H3.2: [I] listened and tried to figure out the type of questions whether it was requesting for help or giving [a] suggestion.
	Students emphasizing on situations	H3.3: [I] focused on the places [ and ] captured on what did the tourists in the [video] clip want. For example, [if it was] the request we must help them. L3.1: [I] focused on the situation.... and captured on tourists' speech. I listened to find out where the situation took place...
	Students emphasizing on vocabulary	L3.3: [I] listened carefully to the vocabulary. [Tourists] spoke too fast [and I] could not catch up.
	Students emphasizing on tourists' gestures from the video clip	L3.1: ... including observed [observing] all the tourists' gestures.

## **2.2 Comprehending strategy** deals with seeking ways to understand the task

Clarifying information	Students clarifying information from the task	H1.3: [I] tried to understand the instruction from the task. For example with the set time [I had to] arrange the answer. [I] emphasized on listing word orders such as "Next, Then [and] After that".
L1 Translating	Students translating L2 to L1	H2.3: ... [I] translated [the instruction] into Thai for some pictures. H3.1: ... [I] translated [the conversation] into Thai to better understand [the situation].

		L3.3: [I] translated into Thai...
Inference	Students using the information from the instruction to understand the tasks	H2.3: [I] used the information from the prompt to understand the task and figured out whether it was “Should do” or “Should not do”...
Analyzing	Students analyzing the information on the task	<p>H1.1: ...[I] analyzed the task on what [I] needed to do. [Figure out whether] it was easy or difficult in order to prepare [for] the information to answer the task.</p> <p>H2.1: [I] ... analyzing the task [whether it was] should do, should not do or just the warning.</p> <p>H3.1: [I] followed up the conversation in the task; and analyzed whether it was complaint or sorry.</p> <p>L1.3: [I] analyzed the instruction and prepared for the information. For example [I] thought of how to use “First, Next and Finally” with the information.</p> <p>L2.3: [I]... analyzed on what the task asked; whether it was “Should” or “Should not do” such as “You should not take a photo because it is a rules”.</p> <p>L3.3: [I] ... analyzed the situation on what the tourists wanted.</p>
Reasoning	Students using the reason to understand the tasks	<p>H2.1: Then, [I] thought about the reason to answer by looking at the picture like in the tourist police, [I said] “If you have problem, please call this number. We can help you”.</p> <p>H3.1: Then, [I] think of the reason to resolve the situation....</p> <p>L1.1: [I] looked for the importance of the place and think of the information and the</p>



		reason. L3.3: ... and [I] thought of the right answer with supporting reasons.
*Imagination	Students imagining their roles as the tour guides to understand the tasks	H1.1: First, we needed to think that we are the real tour guide to understand the task more...  H3.1: We must think that we are the tour guide to better understand the situation...
*Summarizing	Students summarizing the information from the tasks	H1.2: ...[I] summarize [the information] and figure out the picture and the details in order to explain [the picture] uhm [I] really tried to solve the problem in the situation.
<b>2.3 Storing memory</b> strategy deals with finding ways to memorize information to respond the test tasks		
Repeating on what is read or heard	Students repeating the information that they have read or heard from the tasks	H1.2 : [I] ....., reread it [the information] such as which reign was [the place] constructed?, [So that] I could answer the question...  L1.2: [I] tried to reread and remember the information from the pictures and vocabulary.  L2.1: [I] reread the instruction until I could remember everything.  L3.2: ....and [I] reread the instruction.
Associating the existing information	Students linking the learned information from the class with the information on the task to respond the prompt	H3.1: [I] tried to use what I learned in the class to answer [the question] by relating to the information from the tasks.

		<p>L2.1: ... and [I] tried to use the information that I have learned to answer [the task]. I link it with the detail in the task</p> <p>L3.1: [I] used the knowledge from the class to respond.</p>
Summarizing the L2 information	Students summarizing the L2 information to memorize the information to respond the task	L1.3: [I] briefly summarized the English information in order to be easily understood.
Using imagery	Students using imagery to generate, understand or remember the information	<p>H1.1: ... Or [I] memorized the picture to recall the vocabulary and put them into sentences in my own word.</p> <p>H2.3: Most of the time [I] used pictures to memorize. They were very beneficial in memorizing the information and [I] memorized the outstanding features of the picture.</p> <p>L2.3: [I] memorized the picture to give the answer in English. For example, [I said] “You should take off your shoes because it is a rules” or “You should sit polite it is a rules”. [I] mostly used very easy vocabulary that [I] could remember.</p> <p>L3.2: [I] tried to use the picture to memorize. ..</p>
Memorizing linguistic features	Students remembering vocabulary and sentence structures	H1.1: Most of the time, [I] memorized the vocabulary from the class and made sentences from that. ...

	to memorize the information in attempting the test	H3.2: ... [I] memorized the [sentence] structure.
Focusing on main idea of the information	Students emphasizing on the gist of the information to memorize the information	H1.2 : [I] focused on the gist... The example was “The Emerald Buddha Image is [uh][uhm] is symbol of Thai nation[uh] many people go to worship [uh]”. Something like this.
Categorizing information	Students categorizing information to easily memorize the information	L1.1: [I] categorized the piece of information to easily memorize the information such as what was the name?, where was it? , why was it important?
Memorizing situation	Students using situation to memorize the information	H3.3: [I] used the situation to help me memorizing the information.
<b>2.4 Retrieval strategy</b> is used to retrieve L2 linguistic resource and background knowledge to attempt the test tasks		
Transferring L1 linguistic knowledge	Students using L1 grammar to respond the test task	L3.1: [I] sometimes used Thai grammar to construct sentences.
Translating L1 to L2	Students translating L1 to L2 to recall the target language linguistic knowledge in attempting the test task	H2.3: ...and [I] translated Thai into English for some tasks.

		H3.3: [There] might sometimes be Thai into English translation such as the sentences related to incomplete tour program that I need to apologize the tourists. I use what I learned from the class.
Recalling	Students reviewing rereading the answer to recall L2 linguistic knowledge	H1.2: [I] reviewed the answer and reread to recall L2 structures.
Remembering L2 linguistic knowledge	Students memorizing L2 linguistic knowledge including structures and word orders	H1.1: [I] memorized the grammatical structures; for example, “It was built in the reign...”.  H1.3: [I] emphasized on memorizing word orders such as “Next” that was used to explain the next place. H2.1: [I] thought of important L2’s structures. For example, [I] wanted to warn the tourists to keep their wallet [so I] needed to translated some parts in Thai and translated back into L2 as “Please careful for your wallet”.
Using picture to recall L2 linguistic knowledge	Students using picture to recall L2 linguistic knowledge	H2.2: ... and [I] also used the pictures. For example when [I] see people with undershirt picture [I] must warn the tourist to dress politely because it was the important ritual site.

---

**3. Metacognitive strategies deal with the conscious investigation of the test taking process to identify, organize, evaluate and plan the effective ways in attempting the test tasks**

---

**3.1 Goal setting strategy involves identifying the tasks and deciding what to do**

---

Students identifying the information in the tasks and deciding what to do

H1.2: For example, with the picture or tour program [I] needed to understand first and decided on what to do.

H3.2: [I] listened and [tried to] understand [the conversation]. For example, [in the] complaint [situation I] needed to apologize first and then tried to find the supporting reason to respond [the task].

L3.1: [I] found out on what the tourists wanted to prepare the answer.

---

Students identifying the prompt and organizing the information

H1.3: [I] found out what was the requirement from the prompt and selected the information [and] arranged them in my own words.

H2.2: [I] focused on the prompt and prepared the information to respond the prompt. Then, [I] selected the information that was related to each task.

H3.1: [I] focused on what the requirement from the prompt was, thought about what the tourists wanted, tried to find the vocabulary and put them into sentences.

---

Students identifying the situation and deciding what to do

H3.3: [I] primarily focused on the situation to decide how to respond.

L3.2: [I] listened and figured out the situation to answer the question.

---

---

**3.2 Assessment strategy** deals with assessing what is needed, what one has to work with and how well one has done

---

Students assessing what is needed to respond the task

H2.1: [I] thought of technical terms and some sentence structures to use with for example the prohibition.

L1.2: First, [I] assessed on what the prompts asked before attempting it to answer the question.

L2.1: [I] tried to think about the information to respond the task.

L3.2: [I] listened to the conversation in the task and thought about the answer for each task.

---

Students assessing how well they have done on the tasks

H1.1: [I] thought about how well the tourists could understand the answer.

H2.2: After finished the task, [I] assesses how well [I] had done; but, [I] could not change anything. [I was allowed to] record only once.

L1.1: [I] thought about how well I did for each task.

L3.1: [I] thought about how well I did whether I used the right grammar [and] vocabulary with the reason. I wondered how well the tourists could understand my answers.

---

Students thinking of what they have to work with to respond the test tasks

H3.2: [I] figured out what the tourist talked. We needed to find the reason to respond right away. There was no preparation time such as the question about the tourists' attraction.

---

---

**3.3 Planning strategy** involves deciding how to use the existing language knowledge and background knowledge

---

Students deciding to use grammatical sentence structures and background knowledge that were related to prompt

H1.1: [I] thought of how to use the tourism information to respond the prompt? And which structures should be used?

H1.2: Yes, I needed to use tourism knowledge to respond the task. But, if I struggled, I put my opinions especially for the recommendation for the tourists. Then, I think of grammatical structures.

H2.2: [I] used for example with the bus picture [I] thought of how to tell the tourists. [I] need to have tourism knowledge to back up my explanation that they could take the bus. They did have to walk and I used “You should take transfer BTS”. [I] think of structures of the sentences.

L2.2: [I] tried to think of how to use the tourism information to warn the tourists? [I] focused on Thai cultural information on prohibition. Then, [I] made sentences.

---

Students planning to use grammatical sentence structures and background knowledge that were related to situation

H3.1: [I] found the sentence that related to the situation and used the information from the class to respond the prompt by arranging the sentences.

H3.2: [I] figured out what reason to be used to answer the question with the focus on general tourism knowledge to resolve the situation. After that [I] made sentences.

L3.2: [I] selected the tourism information to answer the particular task with appropriate grammatical structures.

L3.3: [I] practiced to be [a] tour guide to resolve the situation. So, [I] must know how to choose the information and speak into English sentences that were appropriate with a particular situation. [I] focused on the tourists' requirement.

---

Students deciding to use vocabulary and background knowledge that were related to prompt

H1.3: [I] selected the most relevant [part] to the prompt for example what was the architectural structure of the Summer Palace? [I] needed to think of the vocabulary.

H2.1: [I] used lots of [background knowledge] especially cultural knowledge particularly for the technical terms. For example "Please be careful" and found the reasons to support the answer related to Thai culture.

H2.3: Sometimes, [I] thought of the technical terms and specific reasons on Thai culture to answer the prompt.

---



## **BIOGRAPHY**

Ms. Malinee Phaiboonnugulkij received her B.A. in English from the Faculty of Arts, Silpakorn University in 2000 and M.A. in Cultural Management from Chulalongkorn University in 2004. Her current position is English instructor at the English Program, the Faculty of Humanities and Social Sciences, Nakhon Ratchasima Rajabhat University. Her research interests include language assessment, LSP instruction and L2 acquisition.