

Tasks in action in Vietnamese EFL high school classrooms: The role
of rehearsal and performance in teaching and learning through oral
tasks

By

NGUYEN THI BAO TRANG

A thesis submitted to Victoria University of Wellington
in fulfilment of the requirements for the degree of Doctor of Philosophy
in Applied Linguistics

Victoria University of Wellington, New Zealand

2013

ACKNOWLEDGEMENTS

This PhD work had been, for me, a process of fruitful discovery and learning. It would not have reached completion without the support of the people whom I thank below.

I owe profound gratitude to my supervisors, Dr Jonathan Newton and Associate Professor David Crabbe not only for their critical minds, thoughtful guidance and insightful feedback but also for their kindness and encouragement throughout these years of study. Without their assistance, this thesis would not have been possible.

I am especially indebted to NZAID for granting me a doctoral scholarship to pursue this study. I wish to thank the NZAID advisers at Victoria University of Wellington, in particular, Inge de Leeuw, Julia Harrison, and Linsell Richards for their kindness and support.

I would also like to express my heartfelt thanks to the Vietnamese teachers and students who participated in my research for their generous time and trust. I sincerely thank the Department of Education and Training, the Head of the school and teaching staff of the school back home in Vietnam for their support and encouragement during my data collection.

I am grateful to the staff at the School of Linguistics and Applied Language Studies (LALS) and Victoria University for creating a supportive and friendly academic environment, which played a nurturing and facilitating role in my research. Special thanks go to Bernie Hambleton, Janet Attrill and Anna Adams for their administrative assistance. I also thank Professor Laurie Bauer, Associate Professor Paul Warren, Dr John Macalister, Dr Stuart Webb, Dr Elaine Vine, Professor Janet Holmes, Dr Jean Parkinson and Dr Angela Joe for the kindness and moral support they offered.

My sincere thanks also go to: Dr Dalice Sim, Statistical Consultant at Victoria for help with statistical analysis; Dr Stuart Webb for letting me audit his Research Methods classes; Tony Quin for assistance with resources and Endnote references; Kirsten Reid for her valuable advice on my oral presentation skills; Dr Xiaodan Gao and again Kirsten Reid for their useful advice on my writing earlier and later respectively in the thesis writing process; Dr Dennis Dawson for helpful advice on graphs and figures; Laila Faisal for useful

suggestions on thesis formatting; Chris Mahoney for allowing me to observe her English Proficiency Programme class, which was my first valuable classroom observation apprenticeship; and Beth Thomas for proofreading my thesis.

I also thank the Faculty of Humanities and Social Sciences at Victoria and NZAID for providing me with grants to carry out data collection in Vietnam and attend conferences.

I am thankful to my varying office mates, Lauren, Dhu, Steve and Jemma, my task-based research group and my LALS PhD group for their research stories. I also thank my Vietnamese friends in Wellington for their friendship and kindness.

Last but not least, I thank my husband, Khiếu, for his mind of science, his own PhD experience and his continued understanding and patience, which in many ways, played a role in my research. I particularly thank my son, Hạo Nhiên and my daughter, Trâm Nhiên for being with me throughout the years. I deeply thank my parents and my parents-in-law for encouraging and supporting me to pursue this PhD study. I dedicate this work to them for their unfailing love during these times and always.

ABSTRACT

Task-based language teaching (TBLT) has attracted considerable attention in research on language teaching and learning. Numerous publications have made a case for TBLT and the role of tasks in learning. TBLT has been introduced in language curricula around the world, including English as a foreign language (EFL) curricula in many countries in Asia. Yet research into tasks in action from both teaching and learning perspectives is rare with scant examination of decisions on task design and implementation that teachers make in the classroom and how their pedagogical decisions are linked to student learning and engagement. The present research addresses these gaps.

The research was conducted in two phases in a Vietnamese high school where a series of task-based EFL textbooks have been adopted to promote curriculum innovation. Phase 1 was a descriptive study which investigated how the Vietnamese EFL teachers implemented oral textbook tasks through adapting task design and creating classroom activity and how learners engaged in the tasks. The data were collected over two and a half months through classroom observations, stimulated recalls and in-depth interviews with teachers and students. The results revealed that the teachers displayed a strong tendency to adapt or replace the textbook tasks, with specific preferences for open over closed tasks, input-independent over input-dependent tasks and divergent over convergent tasks. They also opted for tasks that are not just 'real world', but 'real' to students. Teacher task choices were found to be guided by their own task experimentation, by clearly articulated beliefs about teaching and learning and by a strong orientation to learner engagement.

Decision making by all the teachers reflected a general commitment to a final *public performance* of the task by groups of students. This public performance was preceded by *rehearsal* for the performance, involving students doing the task in pairs or groups to prepare for the performance of the task in front of the class. The terms *rehearsal* and *performance* were used because they captured the teachers' and students' orientation and intent as observed in the lessons and explained in the interviews. Rehearsal and performance constituted two of four identifiable stages of task implementation used by the teachers: pre-task, rehearsal, performance and post-task. Both the teachers and

students valued the notion of performance as a driving force for the use of English and as a social classroom event to engage students in task work. The centrality of public performance in these EFL classrooms, and a lack of empirical evidence about its impact in task-based learning motivated Phase 2 of the thesis.

Phase 2 specifically addressed the impact of task design and learner proficiency on the occurrence and resolution of language-related episodes (LREs) (Swain, 1998) in task rehearsal and on the subsequent take-up in the public performance of the language items which were focussed on in LREs. Three proficiency groups (n=8 dyads in each) from six intact classes carried out two tasks: one problem-solving task (a convergent task) and one debate task (a divergent task), with a 15-minute rehearsal for their performance. The first group was composed of dyad members of the same higher proficiency (HH); the second group consisted of mixed proficiency dyads (HL) and the third group was lower proficiency dyads (LL). The total data included 48 rehearsals and 48 corresponding performances collected in normal classroom hours. Students were also interviewed after they had finished all the tasks.

The results showed that task design and proficiency affected not only the occurrence and resolution of LREs in task rehearsal but also uptake in the public performance.

Specifically, while the problem-solving task induced more LREs, the debate task was more conducive to uptake because the latter task, from the students' perspective, lent itself to performance in ways that the former did not. Overall lower proficiency dyads produced more LREs in rehearsal than higher proficiency dyads. However, it was how LREs were resolved rather than the frequency of LREs that correlated positively with successful uptake in performance. Proficiency also influenced the problem-solving strategies that the learners adopted to prepare for the public performance.

Taken as a whole, this thesis suggests that teacher thinking plays an essential role in transforming tasks in classrooms, and that building in *performance* to tasks and *rehearsal* for that performance may contribute to language learning and development. The research has useful implications for task design and implementation, as well as for theory and research methodology.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	v
TABLE OF CONTENTS.....	vii
LIST OF TABLES	xiv
LIST OF FIGURES.....	xv
ABBREVIATIONS USED IN THE THESIS	xvii
TRANSCRIPTION SYMBOLS USED IN THE THESIS	xviii
Chapter 1 INTRODUCTION	1
1.1 Introduction	1
1.2 Research context	2
1.3 Overview of the research.....	4
1.4 Significance of the research	5
1.5 Organisation of the research.....	6
Chapter 2 LITERATURE REVIEW 1: TASK-BASED LANGUAGE TEACHING, TASKS, FOCUS ON FORM AND CLASSROOM TASK IMPLEMENTATION.....	7
2.1 Introduction	7
2.2 TBLT	7
2.3 Defining tasks	9
2.3.1 Task as learning goal	10
2.3.2 Task as teaching/learning activity.....	12
2.4 Focus on form in TBLT.....	13
2.4.1 A psycholinguistic approach	13
2.4.2 A socio-cultural approach.....	16
2.4.3 A cognitive approach	21
2.4.4 Summary: Focus on form approaches in TBLT	23
2.5 Classroom task implementation.....	24
2.6 Summary and link to the Phase 1 study.....	28

Chapter 3	LITERATURE REVIEW 2: REHEARSAL, PUBLIC PERFORMANCE, LANGUAGE-RELATED EPISODES (LREs) AND L2 LEARNING	31
3.1	Introduction	31
3.2	Pre-task planning.....	31
3.2.1	Strategic planning	32
3.2.2	Rehearsal.....	33
3.3	Public performance	36
3.4	LREs & L2 learning	38
3.4.1	LREs & FFEs.....	38
3.4.2	Focus-on-form episodes (FFE) and L2 learning.....	39
3.4.3	LREs and L2 learning	41
3.4.3.1	LREs and L2 learning measured by tailor-made post-tests.....	42
3.4.3.2	LREs and L2 learning measured by subsequent written performance.....	45
3.4.3.3	LREs and L2 learning measured by subsequent oral performance.....	47
3.5.	Task types and LREs	49
3.5.1	Task types and negotiation of meaning.....	49
3.5.2	Task types and LREs	51
3.6	Proficiency and LREs	54
3.7	The linguistic focus of LREs and L2 learning.....	57
3.8	Summary and link to the Phase 2 study	58
Chapter 4	THE PHASE 1 STUDY: METHODOLOGY.....	61
4.1	Introduction	61
4.2	The research site	62
4.3	The teacher and student participants	62
4.4	Data collection methods.....	65
4.4.1	Classroom observation.....	66
4.4.1.1	Multiple case studies.....	67
4.4.1.2	Video recordings	68
4.4.1.3	Audio recordings	69

4.4.1.4	Field notes and classroom materials.....	69
4.4.2	Stimulated recall and in-depth interviews with the teachers	69
4.4.3	Stimulated recall focus group interviews with students.....	72
4.4.4	Summary of data collection methods for Phase 1.....	73
4.5	Data collection procedures	74
4.5.1	Ethics	74
4.5.2	Piloting	75
4.5.3	Main data collection	76
4.5.4	Summary of data collection procedures	76
4.6	Data analysis.....	77
4.6.1	Analysing the teachers' use of textbook tasks	78
4.6.1.1	How closely the teachers followed the textbooks	78
4.6.1.2	Task design features.....	80
4.6.2	Analysing the teachers' procedures of task implementation.....	88
4.6.3	Analysing the interview data	89
4.6.4	Analysing the learners' task-based interaction data	91
4.6.4.1	The data set	91
4.6.4.2	Coding LREs in rehearsal and uptake in performance.....	91
4.6.4.3	Coding amounts of L1 and L2 use	92
4.7	Validity and reliability	93
4.7.1	Validity	93
4.7.2	Reliability	95
4.8	Summary	95
Chapter 5	THE PHASE 1 STUDY: HOW THE TEACHERS IMPLEMENTED TEXTBOOK TASKS.....	97
5.1	Introduction	97
5.2	How closely did the teachers follow the textbook tasks?.....	97
5.3	In what ways did the design features of the teachers' tasks diverge from the textbook tasks, and why did they diverge from them?	98

5.3.1	In what ways did the design features of the teachers' tasks diverge from the textbook tasks?	98
5.3.2	Why did the teachers diverge from the textbook tasks?	100
5.3.2.1	Open-ended tasks	102
5.3.2.2	Authentic tasks	105
5.4	What task implementation procedures did the teachers use, and why did they use them?...	110
5.4.1	Pre-task	110
5.4.1.1	The teachers' pre-task work	111
5.4.1.2	The teachers' rationales for pre-task work	113
5.4.1.3	Students' perceptions of the learning opportunities offered in the pre-task work	119
5.4.1.4	Pre-task work: Concluding remarks	122
5.4.2	During-task: Rehearsal-performance.....	123
5.4.3	Post-task.....	128
5.4.4	Rehearsal-Performance: Concluding remarks.....	129
5.5	Summary of the chapter	132
Chapter 6	THE PHASE 1 STUDY: HOW THE STUDENTS ENGAGED IN TASKS	133
6.1	Introduction	133
6.2	To what extent did the Vietnamese high school students attend to form while rehearsing for the performance of communicative tasks? If so, how?	133
6.3	To what extent and for what purpose did the students use L1 in task rehearsal?	135
6.4	To what extent did the students use items accurately in performance that had been subject to LREs in rehearsal?.....	141
6.5	How did the students perceive communicative tasks, task rehearsal and performance? ...	143
6.6	Summary of the chapter	148
Chapter 7	THE PHASE 2 STUDY: METHODOLOGY	149
7.1	Introduction	149
7.2	Research questions.....	149
7.3	Student participants	150

7.3.1	Proficiency	150
7.3.2	Gender.....	152
7.4	The tasks	153
7.4.1	Task topic.....	153
7.4.2	Task type	154
7.5	Study design	156
7.6	Data collection procedures	159
7.6.1	Ethics	159
7.6.2	Briefings with the teachers.....	160
7.6.3	The practice sessions	160
7.6.4	Piloting	161
7.6.5	The main data collection	162
7.6.5.1	Rehearsal	162
7.6.5.2	Performance	163
7.6.5.3	Student interviews	164
7.7	Data analysis.....	164
7.7.1	The data set	164
7.7.2	Coding LREs in task rehearsal	164
7.7.3	Coding the linguistic focus of LREs.....	168
7.7.3.1	Lexical LREs	169
7.7.3.2	Grammatical LREs.....	171
7.7.4	Coding resolutions of LREs	173
7.7.4.1	Correctly solved LREs.....	173
7.7.4.2	Incorrectly solved LREs	174
7.7.4.3	Unsolved LREs	175
7.7.5	Coding uptake in task performance.....	177
7.7.5.1	The meaning of uptake in the current study	177
7.7.5.2	The process-product approach	178
7.7.5.3	Coding uptake	179
7.7.6	Inter-coder reliability	188

7.7.7	Statistical analysis	190
7.7.8	Qualitative analysis	191
7.8	Validity and reliability	191
7.8.1	Validity	191
7.8.2	Reliability	192
7.9	Summary	193
Chapter 8	THE PHASE 2 STUDY: LRES IN TASK REHEARSAL	195
8.1	Introduction	195
8.2	Findings	195
8.2.1	Did tasks and proficiency impact on the occurrences of LREs in task rehearsal? ..	195
8.2.2	How did Vietnamese EFL high school learners resolve language problems in task rehearsal?	201
8.2.3	Did task type and proficiency affect how LREs were resolved in task rehearsal? ..	203
8.2.4	Did the linguistic focus of LREs affect how learners resolved them in task rehearsal?	206
8.3	Discussion	212
8.3.1	Rehearsal and task effects on LRE occurrences and resolutions	213
8.3.2	Rehearsal and proficiency effects on LRE occurrences and resolutions	219
8.3.3	Rehearsal and the effects of linguistic focus of LREs on LRE resolutions	223
8.4	Summary	228
Chapter 9	THE PHASE 2 STUDY: UPTAKE IN TASK PERFORMANCE	231
9.1	Introduction	231
9.2	Findings	231
9.2.1	To what extent did learning opportunities operationalised as LREs in rehearsal lead to uptake in performance? Did task type and proficiency impact on the level of uptake?	231
9.2.2	Did task type and proficiency affect different types of uptake in task performance? .	235
9.2.3	What was the relationship between LRE resolutions in task rehearsal and uptake in task performance?	241

9.2.4	Did uptake in task performance differ by linguistic focus?	245
9.3	Discussion.....	252
9.3.1	Task effects on uptake	252
9.3.2	Proficiency effects on uptake	257
9.3.3	Relationship between LRE resolutions and uptake	261
9.3.4	The linguistic focus of LREs and uptake.....	267
9.3.5	Problem-solving strategies: Proficiency and performance pressure	269
9.4	Summary	275
Chapter 10	SUMMARY, IMPLICATIONS AND CONCLUSIONS.....	277
10.1	Introduction	277
10.2	Summary of findings.....	277
10.2.1	Phase 1	277
10.2.2	Phase 2	279
10.3	Pedagogical implications.....	280
10.3.1	Task design	280
10.3.2	Task implementation	282
10.3.2.1	Forms of performance	283
10.3.2.2	Student pairing.....	284
10.3.2.3	Pre-task and post-task work	285
10.3.2.4	Teacher thinking.....	286
10.3.2.5	Language development	287
10.4	Methodological implications.....	288
10.4.1	Research design	288
10.4.2	Data analysis.....	288
10.5	Theoretical implications.....	290
10.6	Limitations and future research directions	291
10.7	Contextual reflection	294
10.8	Personal reflection.....	295

REFERENCES	297
APPENDICES.....	325

LIST OF TABLES

TABLE 2.1: Task definitions.....	9
TABLE 3.1: Willis's (1996) task-based framework (p.52).....	36
TABLE 3.2: Skehan's (1996a) task-based framework (p.54)	37
TABLE 3.3: LREs and language learning by various measures	42
TABLE 3.4: Swain & Lapkin's (2002) study design.....	46
TABLE 4.1: The teacher and class participants	64
TABLE 4.2: Classroom observation scheme.....	67
TABLE 4.3: Summary of data collection methods.....	74
TABLE 4.4: Coding teachers' use of oral tasks from the textbooks.....	78
TABLE 4.5: Features of task design modified by the teachers.....	81
TABLE 5.1: Teacher use of (oral) textbook tasks.....	98
TABLE 5.2: Design features of textbook and teacher tasks	99
TABLE 5.3: The socio-affective flavour in the nine teachers' talking about tasks	101
TABLE 5.4: Teacher action in the pre-task phase.....	112
TABLE 5.5: Locating public performance (adapted from Skehan, 2007b, p.61)	124
TABLE 6.1: Occurrences of LREs.....	133
TABLE 6.2: Amounts of L1 and L2 use by turn in task rehearsal	136
TABLE 6.3: Amounts of L1 and L2 use by word in task rehearsal	136
TABLE 7.1: Proficiency groups and gender	153
TABLE 7.2: Design of the study	157
TABLE 7.3: A sample of reliability data for coding LRE resolutions.....	189
TABLE 7.4: Reliability results.....	190

TABLE 8.1:	LREs by task and proficiency.....	196
TABLE 8.2:	Summary of the RM ANOVA results on LREs across task type, linguistic focus and proficiency	198
TABLE 8.3:	LRE resolutions by task and proficiency.....	201
TABLE 8.4:	Summary of the RM ANOVA results for LRE resolutions	203
TABLE 8.5:	LRE resolution by linguistic focus.....	207
TABLE 8.6:	Summary of the RM ANOVA results for LRE resolutions by linguistic focus and proficiency	209
TABLE 8.7:	Individual data for the HL dyads	222
TABLE 9.1:	Total uptake across tasks and proficiency groups	232
TABLE 9.2:	Summary of the RM ANOVA results for the total uptake in task performance	234
TABLE 9.3:	Uptake types in task performance by task and proficiency	236
TABLE 9.4:	Summary of the RM ANOVA results for uptake types in task performance	238
TABLE 9.5:	Summary of correlation results for LRE resolutions by task	242
TABLE 9.6:	Percentages of unsuccessful uptake out of correctly solved LREs	243
TABLE 9.7:	Percentages of successful uptake out of correctly resolved LREs	244
TABLE 9.8:	Percentages of incorrect uptake out of incorrectly solved LREs	245
TABLE 9.9:	Uptake types by linguistic focus across proficiency groups	246
TABLE 9.10:	Summary of the RM ANOVA results for uptake types across linguistic focus and proficiency	248
TABLE 10.1:	Teacher use and implementation of textbook tasks	278

LIST OF FIGURES

FIGURE 4.1:	Summary of data collection procedures	77
FIGURE 5.1:	Design features of textbook and teacher tasks	100
FIGURE 5.2:	Task implementation procedures by the teachers	110
FIGURE 7.1:	Procedures for selecting proficiency groups	152
FIGURE 7.2:	The scheme for coding uptake in task performance	181

FIGURE 8.1:	LREs cross tasks and proficiency groups	197
FIGURE 8.2:	LREs by task, proficiency and linguistic focus.....	200
FIGURE 8.3:	LRE resolutions across tasks and proficiency groups.....	202
FIGURE 8.4:	Correctly solved LREs by task and proficiency.....	204
FIGURE 8.5:	Incorrectly solved LREs by task and proficiency	205
FIGURE 8.6:	Unsolved LREs by task and proficiency	206
FIGURE 8.7:	LRE resolution by linguistic focus.....	208
FIGURE 8.8:	Correctly solved LREs by linguistic focus.....	210
FIGURE 8.9:	Incorrectly solved LREs by linguistic focus	211
FIGURE 8.10:	Unsolved LREs by linguistic focus.....	212
FIGURE 9.1:	Total uptake by task and proficiency	233
FIGURE 9.2:	Total uptake across tasks and proficiency groups	235
FIGURE 9.3:	Uptake types across tasks and proficiency groups	237
FIGURE 9.4:	Successful uptake across tasks and proficiency groups	239
FIGURE 9.5:	Unsuccessful uptake across tasks and proficiency groups.....	240
FIGURE 9.6:	Incorrect uptake across tasks and proficiency groups.....	241
FIGURE 9.7:	Uptake by linguistic focus across proficiency groups	247
FIGURE 9.8:	Successful uptake by linguistic focus across proficiency groups	249
FIGURE 9.9:	Unsuccessful uptake by linguistic focus across proficiency groups.....	250
FIGURE 9.10:	Incorrect uptake by linguistic focus across proficiency groups	250
FIGURE 9.11:	No uptake by linguistic focus across proficiency groups.....	251

ABBREVIATIONS USED IN THE THESIS

EFL	English as a foreign language
ESL	English as a second language
L1	First language
L2	The target language (second, third or foreign)
TBLT	Task-based language teaching
CLT	Communicative language teaching
LRE	Language-related episode
FFE	Focus-on-form episode
PPP	Presentation-Practice-Production
ZPD	Zone of proximal development
RM ANOVA	Repeated Measures Analysis of Variance

TRANSCRIPTION SYMBOLS USED IN THE THESIS

Symbol	Meaning
()	English translation of Vietnamese L1 utterances
[]	The researcher's comments/explanations
(.)	Short pause (.05-3 seconds)
...	Long pause (>= 4 seconds)
(...)	Deleted texts
[Overlapping
xxx	Indiscernible texts
-(ter-)	Incomplete pronunciation of a word
a-b-c-d-e-f	The speaker is spelling out the word.
?	The speaker is asking a question or raising the intonation.
.	Falling intonation
,	Continuing intonation
“ ”	The speaker is reading out from task input provided.
Upper case	Word emphasis
(e.g., VOLUNTEER)	
!	Emotional emphasis at the end of a phrase/sentence
/ /	Phonetic transcription
T	Teacher
S	Student
S1	Student 1
S2	Student 2
Ss	Many students
R	Researcher

Chapter 1 INTRODUCTION

1.1 Introduction

In recent decades, task-based language teaching (TBLT) has received growing attention from both researchers and practitioners. TBLT locates L2 teaching and learning in tasks, that is in meaning-focussed activities where learners use whatever language resources they have to carry out the task to achieve its *non-linguistic* outcome (Ellis, 2009a). This emphasis on a non-linguistic outcome distinguishes tasks from exercises which focus on accurate usage of pre-selected language items (Ellis, 2003). TBLT posits that learners learn the language through transacting tasks that trigger ‘holistic language use’ (Samuda & Bygate, 2008) or similar cognitive processes as required of real life communication (Ellis, 2003, 2010a; Long & Crookes, 1992; Skehan, 1998; Van den Branden, 2006a; Willis & Willis, 2007). In essence, TBLT places emphasis on putting language to use to achieve a task outcome rather than on focusing on language forms for their own sake.

TBLT has become increasingly popular in language teaching around the world, including Asia (Butler, 2011; Nunan, 2003). It has been adopted as EFL curriculum innovation in many Asian countries such as in China (Deng & Carless, 2009; Zhang, 2007), in Hong Kong (Carless, 2002, 2003, 2004, 2007, 2008), in South Korea (Jeon & Hahn, 2006), in Thailand (McDonough & Chaikitmongkol, 2007) and in Vietnam (Barnard & Nguyen, 2010; Le & Barnard, 2009).

Research has thus begun to investigate the implementation of TBLT in these Asian EFL contexts. This body of research has, in the main, identified factors in contributing to the limited uptake of TBLT in these settings (see Butler, 2011 for a recent review). Less research has focussed on the analysis of teacher tasks in action and the underlying teacher thinking *vis-à-vis* prescribed textbook tasks in terms of task design features and methodology (pre-task, during-task, post-task). Similarly, little has been reported on the relationship between teacher thinking underlying the way they design and implement tasks, and student engagement in tasks and learning outcomes, an important relationship given the tandem nature of teaching and learning.

Given the meaning-focussed nature of tasks, the learning dimension in tasks has been a concern of both practitioners and task researchers (Bygate & Samuda, 2009). Teachers are often worried about whether students learn anything during task-based interaction in EFL contexts in which students share a first language (L1) (e.g., McDonough, 2004). Similarly, task researchers are concerned that learners might resort to communicative strategies to successfully complete a given task, without 'pressurizing' their language use (Bygate, 1996; Bygate & Samuda, 2005, 2009; Skehan, 1998, 2007a; Skehan & Foster, 2001; Skehan, Xiaoyue, Qian, & Wang, 2012). These researchers argue that learners need to be 'pushed' to *communicate for learning* (Bygate & Samuda, 2009). Research has sought to manipulate task design features and task conditions to achieve this, as seen, for example, in the numerous studies on pre-task planning and focus on form in the during-task stage. In contrast, although public performance has long been recommended in task-based frameworks (Skehan, 1996a, 1998; Willis, 1996), and it might be commonly practised in classrooms, it has been rarely researched (cf. Skehan & Foster, 1997).

The current research addressed these gaps.

1.2 Research context

The foreign languages being taught in Vietnam have followed the political, economic and socio-cultural statuses of those foreign languages (Baecher & Dang, 2011; Phan, 2009). Today English has taken the predominant position as a foreign language in Vietnam, replacing Chinese, French and Russian. Teaching and learning EFL in Vietnam has burgeoned, since Vietnam's *Đổi mới* (Renovation) policy in 1986 and accelerated when the country officially joined the World Trade Organisation (WTO) in 2007. With the new role that English now plays in the development strategy of Vietnam, education authorities in Vietnam have called for innovative EFL instruction with the aim of producing a labour force with English proficiency to facilitate the country's modernisation, industrialisation and global integration.

The Vietnamese Ministry of Education and Training has specifically emphasised the importance of the development of communicative competence for students via

innovative teaching methodology that provides students with opportunities to use the target language. Methodological innovation in EFL instruction has been called for not only at university but also at primary, secondary and high school levels in Vietnam.

The national English curriculum for Vietnamese high school students has been renewed through a series of new textbooks which were officially approved and adopted in 2006 (Le & Barnard, 2009; MOET, 2010). The new textbooks, according to their authors, reflect a communicative approach, learner-centeredness and *task-based teaching* as central focus (Hoang et al., 2006, 2007). The textbooks emphasise (1) “tasks as main activities to develop learners’ communicative competence”; (2) learners as “proactive and creative agents in the learning process”; and (3) teachers as “organiser, monitor, mediator, consultant, participant, and knowledge provider” (Hoang et al., 2007, p.6, translated from Vietnamese). This curriculum is by now well embedded into high school English language teaching in Vietnam.

Although the new curriculum has been well in use for a while, research into its implementation is rare, with only two small-scale studies to date (Le & Barnard, 2009; Barnard & Nguyen, 2010). The former study looked at teacher classroom practice and found the dominance of teacher-fronted grammar-based teaching. The latter study focussed on teacher attitudes towards TBLT by means of ‘narrative frames’ which were guided reflections that teachers were asked to write. It found that teachers reported appreciating the values of communicative tasks, but emphasised the importance of explicit teaching of grammar, vocabulary and pronunciation. The paucity of research within the context of the new task-based programme for Vietnamese high school students is unfortunate, given the curriculum targeting millions of high school teachers and students in Vietnam. More empirical evidence on tasks in action is obviously needed, particularly from both teacher and student perspectives, and from both teaching and learning.

Furthermore, existing research into general EFL instruction in Vietnam, mainly in university settings, has shown a lack of meaningful communication in the classroom (Bock, 2000; Evans, 1999, cited in Nguyen, 2003; Le, 2001; Tomlinson & Bao, 2004). Evans (1999) claimed that most classroom learning in Vietnamese EFL classrooms is

“passive” and “receptive” (p.52). Tomlinson and Bao (2004) also found that “spontaneous discourse was rare” (p.99). Le (2001) further described the Vietnamese EFL classroom as “a cultural island where the teacher is expected to be the sole provider of experience in the target language” and learners “are expected to sit in silence unless the teacher calls them individually to speak” (pp.35-36). However, such claims need to be tested in other Vietnamese EFL contexts, especially in high school classrooms, an underrepresented context in EFL research in Vietnam and elsewhere.

1.3 Overview of the research

The current research is an answer to the pressing call for a ‘researched pedagogy’ by investigating tasks in classrooms (Bygate, 2011; Bygate, Norris, & Van den Branden, 2009). It took place in EFL classrooms in a public high school in Vietnam which used the new task-based textbooks. It set out to achieve four main aims.

First, the research sought to investigate a neglected area in TBLT, the ways teachers implement textbook tasks, and the teacher thinking that lies behind their implementation choices.

Second, it sought to document an empirical link between the choices teachers make in implementing tasks and student task engagement and perceptions. This area has rarely been studied in TBLT research.

Third, the current research sought to examine a particular feature of task implementation, *public performance* (thereafter *performance*) and preparation for that performance (thereafter *rehearsal*) in this teaching context and one that has not been much considered in TBLT research, especially from both teacher and student perspectives. The terms *rehearsal* and *performance* were used in the thesis because they captured the teachers’ and students’ orientation and intent as observed in the lessons and explained in the interviews.

Fourth, it sought to further examine the effects of task types and learner proficiency on the extent to which language-related episodes (LREs) that arose in dyadic task rehearsal were taken up in the dyadic (public) performance of the same task. Although previous

research has shown task type (e.g., Alegría de la Colina & García Mayo, 2007; García Mayo, 2002; Storch, 2001; Swain & Lapkin, 2001) and proficiency (e.g., Kim & McDonough, 2008; Leaser, 2004; Watanabe & Swain, 2007; Williams, 1999, 2001) have major effects on the frequency and types (grammatical/lexical) of LREs, whether and how these two variables impact on L2 learning by means of *rehearsal-performance*, has not been explored.

The current research focuses on *oral* tasks for four reasons: (1) they are a central component in the new curriculum; (2) research has shown meaningful communication in Vietnamese EFL classrooms, especially in high school settings, is infrequent (see 1.2 above); (3) oral pair/group tasks have been reported as difficult to successfully implement in Asia due to contextual constraints such as L1 use, discipline, and paper-and-pencil examinations (e.g., Carless, 2007, 2008; McDonough, 2004; Pham, 2007) (see Butler, 2011 for a recent review); and (4) oral pair/groupwork has been said to be imported from the West, and thus questioned in Asian contexts in which it may conflict with cultural norms (e.g., Ellis, 1996; Holliday, 1997; Littlewood, 2007).

The research was carried out in two phases. Phase 1 was a descriptive study which identified how the teachers enacted oral communicative textbook tasks in their daily classrooms and how learners engaged in the classroom tasks.

Phase 2 was pedagogically motivated by the findings from Phase 1, with a detailed focus on (public) *performance* and *rehearsal* for that performance, the centrality of these EFL classrooms. It was a mixed design quasi-experimental study that investigated the effects of task types and proficiency on (1) the occurrences of LREs, types of LREs (lexical/grammatical) and how LREs were resolved in dyadic task rehearsal; and (2) the take-up of the LRE-specific language items in dyadic task performance.

1.4 Significance of the research

The current research is of significance in several ways. First, personally the research has value to me as a practitioner using TBLT in my teaching. I had been a high school teacher of English for 13 years in Vietnam before taking PhD study leave, and had used the

mandated new task-based English textbooks for five years. So by conducting this research, I am acquiring a 'researched pedagogy' for my own EFL teaching.

Second, given the current importance of EFL instruction in Vietnam to meet its demands for proficient users of English to serve its cause of global integration, the findings from this research provide grounds for potential enhancement of EFL teaching and learning in Vietnam. They provide important insights into teacher task choices and learner engagement in different types of tasks and especially in the rehearsal- performance approach to tasks taken by the teachers in the study. These insights can be applicable to the work of millions of EFL high school teachers and students in 64 provinces and cities in Vietnam who are using the new task-based textbooks. It is also hoped that the findings can be applied beyond Vietnam, in teaching contexts in which tasks are used.

Finally, the present research provides emic perspectives on TBLT from teachers and students, perspectives that can inform the work of textbook designers, teacher educators, and policy makers, as well as task theorists and researchers.

1.5 Organisation of the research

The thesis consists of ten chapters. This first chapter has introduced the thesis. Chapter 2 and Chapter 3 deal with the relevant literature reviews. Chapter 4 presents the methodology for Phase 1, followed by Chapter 5 and Chapter 6, which report and discuss its results concerning the teachers using and implementing textbook tasks and how the students engaged in classroom tasks respectively. Chapter 7 presents the methodology for Phase 2. Chapter 8 and Chapter 9 report and discuss its findings with respect to the effects of tasks and proficiency on the occurrence of LREs, and LRE resolutions in task rehearsal and uptake in task performance. Chapter 10 is the conclusion which summarises the main findings of the thesis, discussing its pedagogical, methodological and theoretical implications. It also discusses the thesis's limitations and proposes future research directions. It concludes with my contextual and personal reflection.

In the next two chapters I will review the relevant literature to the present research.

Chapter 2 LITERATURE REVIEW 1: TASK-BASED LANGUAGE TEACHING, TASKS, FOCUS ON FORM AND CLASSROOM TASK IMPLEMENTATION

2.1 Introduction

This chapter first briefly describes and discusses what task-based language teaching (TBLT) entails, defining tasks and a central tenet in TBLT, focus on form. It then reviews studies on tasks in action and teacher thinking and discusses limitations of previous research that provided the impetus for Phase 1 of the current research.

2.2 TBLT

TBLT was first proposed in the 1980s in seminal papers by Breen and Candlin (1980), Long (1985a) and Prabhu (1987) in response to dissatisfaction with traditional language teaching approaches that involve teaching pre-selected language items in discrete isolated blocks (Long & Crookes, 1992; Long & Norris, 2009; Van den Branden, 2006a). These approaches include the structural approach and the functional-notional approach (weak communicative language teaching, CLT) that separate structures, notions and functions as units of analysis to teach and expect the learner to *synthesise* them as needed for communicative purposes, and thus adopt 'synthetic' syllabi (Wilkins, 1976)¹ or Type A syllabuses (White, 1988), categorised as focus on forms approaches by Long (1991) and Long and Robinson (1998). Long (2007) strongly argues that "focus on forms attempts the impossible: to impose a pre-set, external linguistic syllabus on learners, riding roughshod over individual differences in readiness to learn, even within classes of students with the same overall "proficiency". It is psycholinguistically untenable"(p.121).

¹ While Wilkins (1976) puts the functional-notional approach on the analytic type, Long and Crookes (1992) argue that language functions such as requesting, apologizing, and inviting are linguistic units, and that synthetic syllabuses also include *covert* units of analysis such as topic and situation since they are often seeded with the 'structure of the day'.

The common presentation-practice- production (PPP)² paradigm has also been criticised for similar reasons, being (1) incompatible with the learner's 'internal syllabus'; and (2) unrealistic for failing to cater for individualised learning. Skehan (2002) elaborates:

Classes are made up of a range of individuals with different talents, styles, and motivations. As a result, what may be presented to a group may only be appropriate for a small number of learners within that group. For others it may be too difficult, for yet others too easy. In the second case, it is an ineffective but excusable waste of time. In the first case, where the material is too difficult, the time spent focusing on the language element may also be a waste of time, but in this case it will leave the particular language point untouched, since the learner isn't ready to absorb it, although the teacher and class may assume that the point has been learnt. This is altogether more serious, since instruction continues, things may get worse because the initial learning couldn't occur (p.290).

Willis (1996, p.134) also takes three issues with the PPP paradigm: (1) learners at the final P might not use the language items presented; (2) if learners are required to use pre-selected linguistic items they cannot produce the language freely. In this way intervention such as making use of targeted items compulsory would convert tasks into exercises (Ellis, 2003); (3) there are cases when students overuse language items presented in advance, leading to mechanical artificial use. As a result, despite years of formal instruction, many learners cannot communicate in real life situations (cf. DeKeyser, 2001; Klapper, 2003).

In contrast, TBLT does not take preselected language items as the starting point for teachers to teach and for students to master one by one in accumulation, but posits that "the performance of functional tasks involving meaningful language use is the starting point, primary mechanism, and final goal of educational activity"(Van den Branden, Bygate, & Norris, 2009, p.6). The centrality is that "people learn a language not only *in order to* use the target language for functional purposes, but also *by doing so*" (Van den Branden, 2012, p.133, original italics). TBLT, in the words of Skehan (1996b), "makes the assumption that transacting tasks ... will engage naturalistic acquisitional mechanisms,

² Learners are first presented with a particular language form. Then they practice this form in a controlled manner to display accuracy. Finally they are given opportunities to use it in a 'free' situation (Willis, 1996, p. 134).

cause the underlying interlanguage system to be stretched, and drive development forward” (p.95).

But what is a task in TBLT?

2.3 Defining tasks

Tasks have been defined in various ways in the literature (see Bygate, Skehan, & Swain, 2001; Ellis, 2003, 2009a; Samuda & Bygate, 2008; Skehan, 1998; Van den Branden, 2006a; Willis, 2004). Table 2.1 provides some selected task definitions.

TABLE 2.1: Task definitions

Author	Task definition
Long (1985a)	A piece of work undertaken for oneself or for others freely or for some reward. Thus examples of tasks include painting a fence, dressing a child, filling out a form In other words, by task is meant the hundred and one things people do in everyday life, at work, at play, and in between. ‘Tasks’ are the things people will tell you they do if you ask them and they are not applied linguistics. (p.89)
Skehan (1998)	A task is an activity in which -Meaning is primary. -There is some communication problem to solve. -There is some sort of relationship to comparable real world activities. -Task completion has some priority. -The assessment of the task is in terms of outcome. (p.95)
Van den Branden (2006a)	A task is an activity in which a person engages in order to attain an objective, and which necessitates the use of language. (p.4)
Samuda and Bygate (2008)	A task is a holistic activity which engages language use in order to achieve some non-linguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both. (p.69)
Ellis (2009a)	A task is an activity which - Focuses primarily on meaning. - Creates a need for meaning to be made. - Allows learners their own choice of linguistic and non-linguistic means to

complete the task.

- Has a clearly defined outcome to be achieved by means of language, not language being an end in its own right. (p.223)

Overall, common to these task definitions is that task is a meaning-focussed activity with a 'non-linguistic outcome' which students can use whatever linguistic means to complete. Task-based learning is therefore 'holistic', 'meaning-focussed' and 'learner-driven' (Van den Branden et al., 2009, pp.2-3). TBLT hence presents "the target language whole chunks at a time, without linguistic interference or control" and the role of the learner is to *analyse* and induce rules from input and form-meaning mapping in the course of task performance (Long & Crookes, 1992, p.29), and thus belongs to 'analytic' (Wilkins, 1976) or Type B (White, 1988) syllabuses.

The subsequent sections further discuss two key dimensions of task: task as learning goal and task as teaching/learning activity (also see Van den Branden, 2006a).

2.3.1 Task as learning goal

Task as learning goal reflects different views on task authenticity, an aspect that has caused issues with task definition (Skehan, 2003) and that has been subject to criticisms (e.g., Widdowson, 2003). Long's task definition in Table 2.1 above reveals his strong advocacy of the functional use or 'situational authenticity' of tasks, that is, tasks involving a real life situation (Ellis, 2003). Long (2005, 2007; Long & Crookes, 1992) strongly argues that the selection and design of pedagogic tasks should be based on needs analysis with regards to "the real world target tasks learners are preparing to undertake" (p.44). However, the precise link between target tasks and classroom tasks is "vague" (Van den Branden, 2006a, p.6) or "not simple" (Pica, Kanagy, & Falodun, 1993). Furthermore, although needs analysis is "desirable", it is "difficult to obtain" (Skehan, 1996a, p.39). This seems to be true in EFL contexts where relating to target tasks as learning goals might not be perceived in the same ways as in English as a second language (ESL) contexts.

Although task defined by Skehan (Table 2.1 above) (also see Ellis, 2003) also addresses 'the real world resemblance', these researchers take a less stringent view on the issue of authenticity. Ellis (2003) particularly argues that tasks with both situational and interactional authenticity³ should be used because they trigger similar acquisition processes inherent in real world communication. Skehan (2003) re-emphasises that 'real world resemblance' means "the nature of the *response* by the learner ... rather than a form of authenticity, defined only in relation to the real world occurrence of an activity"(Skehan, 2003, p.3, italics added) (also see Skehan, 2007b).

The issue of "how far classroom tasks mirror the real world", according to Willis and Willis (2007) is manifested at three levels: the level of *meaning*, the level of *discourse* and the level of *activity* (p.136). The level of activity corresponds to situational authenticity. Some tasks might not "offer a precise reflection of the real world" or they are 'artificial tasks', "but they do oblige learners to engage in real world meaning and real world discourse" (Willis & Willis, 2007, p.142). Ellis's (2009a) recent task definition clarifies that the 'real world resemblance' aspect means learners can use whatever linguistic and non-linguistic resources to complete the task (see Table 2.1 above) (also see Willis, 1996) and once again stresses that *all* tasks should achieve interactional authenticity. Therefore, it does not necessarily mean that students must achieve the goal of doing the target task in real life. Rather, by doing tasks, they engage in similar processes as required of real life communication and meaning-making, and this is the ultimate goal (Ellis, 2003; Long & Crookes, 1992; Prabhu, 1987; Richards & Rodgers, 2001; Willis, 1996; Willis & Willis, 2007). This contention is reflected in later task definitions in Table 2.1, which emphasise learner "engagement with meaningful language use" (Skehan, 2007b, p.291) to achieve an outcome. Viewing tasks in this way provides legitimate motivations for TBLT in a variety of contexts, not only contexts in which tasks bear direct relevance to the real world activity. "A task is not an action carried out *on* task participants; rather, a task is an activity which participants, themselves, must carry out" (Pica et al., 1993, p.12). In this regard, task authenticity

³ Nunan (2004) used rehearsal and activation tasks to refer to tasks with situational and interactional authenticity respectively.

broadly corresponds to “whether or not students are ‘engaged’ by the task” (Guariento & Morley, 2001, p.350).

2.3.2 Task as teaching/learning activity

Tasks are defined as both the learning goal and learning activity, by which to achieve the goal. Task as teaching/learning activity that shapes the syllabus forms TBLT. However, TBLT is not a ‘monolithic’ approach and has strong and weak forms (Ellis, 2003, 2009a; Skehan, 1996b). The former is synonymous with TBLT while the latter is a kind of task-supported language teaching where tasks are used along with other traditional teaching approaches (Ellis, 2003; Samuda & Bygate, 2008) or in the final P of the PPP paradigm (Skehan, 1996b). According to Samuda and Bygate (2008), this distinction is useful because it underscores the flexible roles of tasks as a pedagogic tool while arguing that each version deserves its own understanding. Van den Branden et al. (2009, p.9) further distinguish two ‘arms’ of TBLT: syllabus specification and teaching procedures. In the former, the content of the syllabus is target tasks as argued by Long and Crookes (1992). In the latter,

Teachers use tasks as the fundamental reference point for their own teaching. This of course they can do even in the context of linguistically defined syllabus (cf. Samuda, 2001), the teaching methodology being task-based even if the syllabus is not (Van den Branden et al., 2009, p.9).⁴

Regarding task as both learning activity and learning goal, some researchers (e.g., Nunan, 1989) argue that the distinction between syllabus and methodology in TBLT is redundant. However, other researchers (Skehan, 1998; Ellis, 2003, 2009) argue that methodology, that is, how a task is implemented (e.g., pre-task, during-task and post-task) can impact task performance. This is also my view on tasks and task implementation throughout the thesis.

⁴ Ellis (2003) distinguishes unfocussed tasks and focussed tasks. The former do not target any language structures while the latter do, although learners are not told to use the targeted items.

2.4 Focus on form in TBLT

One key task feature is the primacy of meaning. However, it is this feature that has led to the criticism that TBLT does not provide adequate opportunities to learn the formal features of language (see Ellis, 2009a). Indeed, one central principle of TBLT is that while focusing primarily on meaning and ‘holistic language use’, learners should pay attention to form (grammar, vocabulary, pronunciation or discourse) or *focus on form* (Ellis, 2003, 2005, 2009a; Long, 1991, 1996; Mackey, 2012; Skehan, 2007b; Swain, 2005).

The term ‘focus on form’ was originally coined by Long (1991) to refer to instruction that “overtly draws students’ attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication” (pp.45-46). This focus on form (FonF), according to Long (1991, 1996; Long & Robinson, 1998), aims to achieve a balance between the two extreme stances: the naturalist approach (Krashen, 1981), and the traditional approach, focus on forms.

Over the years, various approaches to focus on form have differed from the original meanings (see Ellis, 2009a; Ellis, Basturkmen, & Loewen, 2001a, 2001b; Williams, 2005). Yet these approaches share the view that tasks alone are not enough and thus seek various ways to ‘stretch’ and ‘push’ language development forward (Long, 1996, 2007; Skehan, 1998, 2009; Swain, 2005; Willis, 1996). They include a psycholinguistic approach, a socio-cultural approach and a cognitive approach (also see Ellis, 2000; Skehan, 2003, 2007b).

2.4.1 A psycholinguistic approach

A psycholinguistic perspective on focus on form in TBLT draws on the related roles of input, interaction, noticing and output.

Motivated by the input hypothesis (Krashen, 1981, 1982, 1985) which posits that $i+1$ input or input that is slightly beyond the learner’s current level is necessary and sufficient for acquisition to occur, Long (1983, 1985a) argues that learners can obtain comprehensible input through negotiation of meaning (e.g., comprehension checks, clarification requests, confirmation checks) to resolve communication problems. This

forms the basic argument of Long's (1983) early interaction hypothesis that extended the original input hypothesis (Mitchell & Myles, 2004). Long stresses the role of interactional modifications or 'interactive input' (Ellis, 2008) in conversations to improve the comprehensibility of input rather than pre-modified input or 'non-interactive input' (Ellis, 2008).

Research in response to the early interaction hypothesis attempted to show an indirect relationship between negotiation of meaning and acquisition via comprehension (Ellis, 2008). Ellis points out that comprehension involves semantic processing and thus "does not necessitate close attention to linguistic form" (p.251) and that the types of comprehension processes needed for acquisition to take place are not specified. Many researchers (e.g., Faerch & Kasper, 1986) argue that how comprehension facilitates acquisition depends on learners noticing the input and making comparisons between this input and their interlanguage. Long (1996) later specified mechanisms whereby negotiation of meaning can assist acquisition: "communication trouble ... can lead learners to recognize that a linguistic problem exists, switch their attentional focus from message to form, identify the problem and notice the needed item in the input" (p.425).

Research has then focussed on seeking tasks that are conducive to negotiation of meaning such as two-way required information gap tasks, and closed convergent tasks (Duff, 1986; Pica et al., 1993)(see Ellis, 2000, 2003; Mackey, 2012 for reviews). The role of feedback and the way learners modify their output in order to be more comprehensible are also components of focus on form (e.g., Foster & Ohta, 2005; Gass & Mackey, 2007; Long, 1996; Mackey, 2012; Pica, 1994; Shehadeh, 2002, 2004). Long (1996) argues that negotiation of meaning leads to acquisition because "it connects input, internal learner capacities, particularly selective attention, and output in productive ways" (pp.451-452).

The role of producing language in language acquisition was highlighted in the output hypothesis (Swain, 1985, 1995, 2005). This hypothesis challenged the view of comprehensible input as "the only true cause of second language acquisition" (Krashen, 1984, p.61). While acknowledging comprehensible input is necessary, Swain (1985) argues that comprehensible input is not sufficient for language acquisition to occur.

Drawing on her research in French immersion programmes in Canada, she found that, despite exposure to abundant comprehensible input, the French immersion students in her study failed to use the target language accurately and appropriately for two reasons: students did not have sufficient opportunities to use the target language and they were not 'pushed' in their output because they were not pressurised to "be more comprehensible than they already are" (p.249). In Swain's (1985) words:

Conversational exchanges ... are not themselves the source of acquisition derived from comprehensible input. Rather they are the source of acquisition derived from comprehensible output: output that *extends the linguistic repertoire* of the learner as he or she attempts to create precisely and appropriately the meaning desired (p.252, italics added).

Swain (2005) proposes three main functions of output. First, while producing output, learners are confronted with tensions in finding language resources to express what they want to communicate and so they may be directed to notice the 'hole' between

what they *want* to say and what they *can* say, leading them to recognize what they don't know, or know partially..... This may trigger cognitive processes which might generate linguistic knowledge that is new for learners, or which consolidate their existing knowledge. (Swain, 1995, p.126, original emphasis)

This noticing function echoes the noticing hypothesis (Schmidt, 1990, 1994, 2001), which argues that noticing, or 'conscious attention' is crucial in L2 learning. Second, output has the hypothesis testing function in that "output may sometimes be, from the learners' perspective, a "trial run" reflecting their hypothesis of how to say (or to write) their intent" (Swain, 2005, p.476). Third, the metalinguistic/reflective function claims that "using language to reflect on language produced by others or the self, mediates second language learning" (Swain, 2005, p.478). This third function of output is related to socio-cultural theory that will be elaborated later.

Above all, underlying the functions of output is the role of *consciousness* in acquisition. While producing output, learners need to process language syntactically rather than semantically (Ellis, 2008; Swain & Lapkin, 1995). Therefore, output engages learners in deeper language processes than input (Gass, 1997; Skehan, 1998; Swain, 2000).

Producing output also enhances automatisisation and fluency (de Bot, 1996; Skehan, 1998; Swain, 1995), allows learners to bring in their personal voices (Skehan, 1998; Ellis, 2008), and provides ‘auto input’, input from one’s own production (Ellis, 2008).

Empirical research supports the claims in the output hypothesis by showing how opportunities to produce language lead to acquisition (de la Fuente, 2002; Ellis & He, 1999; Izumi, 2003; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Nobuyoshi & Ellis, 1993). Research has also focussed on the role of feedback (see Mackey, 2007, 2012 for recent reviews), again providing evidence in support of the output hypothesis. Drawing on the roles of input, interaction, and output, the interaction hypothesis now provides richer insights into how negotiation can help language learning through both positive and negative evidence (Ellis, 2008, p.255). The interaction hypothesis has been recently referred to as ‘the interaction approach’ (Gass & Mackey, 2007; Mackey, 2012; Mackey, Abbuhl, & Gass, 2012) which argues that the question is now not whether interaction influences learning, but rather how it affords opportunities for learning (Mackey, 2012).

Other research studies by Swain and her colleagues have focussed on output as a cognitive process that mediates learning. One focus of this research is language-related episodes (LREs) that arise during *student* collaborative work and evidence of learning (e.g., Brooks & Swain, 2009; Donato, 1994; Storch, 2002a, 2002b; Swain, 1998; Swain & Lapkin, 2001). This body of research has investigated tasks that encourage students to discuss and resolve LREs during collaborative talk, and in so doing to increase awareness of language forms, leading to internalisation of these forms or consolidation of existing language knowledge (Swain & Lapkin, 1995)(also see Chapter 3). As Skehan (2007b) points out, “the focus, in this case, is psycholinguistics, and how form is brought into focus, but the means connect with socio-cultural theory” (p.295). A socio-cultural perspective on focus on form is addressed next.

2.4.2 A socio-cultural approach

A socio-cultural perspective centres on ‘mediated learning’ (Lantolf, 2000). Socio-cultural theory, drawing on the work of Vygotsky (1978), among others, proposes that

human cognition development is mediated by means of social interaction with others, self and artefacts. A socio-cultural approach to focus on form emphasises the role of scaffolding, especially the zone of proximal development (ZPD), the role of language as a cognitive tool that mediates learning, and thus the role of L1 and learner agency.

First, I look at the zone of proximal development (ZPD). From a Vygotskian perspective, human cognition is mediated via social interaction. It develops and evolves first and most importantly interpsychologically through interaction between people (Vygotsky, 1978, 1987; Wertsch, 1985). The ZPD was originally defined as the level of development that one can attain with assistance which otherwise cannot be achieved without being guided and assisted (Vygotsky, 1978, 1987). The ZPD requires social interaction especially with a more capable interlocutor such as a teacher, an adult or a more proficient learner. From a second language learning perspective, Ohta (2001) re-defined ZPD as “the distance between the actual developmental level as determined by individual linguistic production, and the level of potential development as determined through language produced collaboratively with a teacher or peer” (p.9).

Learner-learner interaction studies in SLA have progressed beyond expert-novice interaction to show that learners can benefit from peer-peer interaction with each other (e.g., Donato, 1994; Ohta, 2000, 2001). According to Lantolf (2000), the ZPD should be “more appropriately conceived of as the collaborative construction of opportunities” (p.17). In this way, the ZPD is different from the $i+1$ input in Krashen’s input hypothesis in that the former emphasises the role of the learner as agent, and co-construction while the latter the role of input (also see Ellis, 2003, 2008; Mitchell & Myles, 2004).

Second, I look at the use of language as a cognitive tool or a “tool for thought” (Mitchell & Myles, 2004, p.194). This view relates to the metalinguistic or reflective function of output (Swain, 2005). Swain noted that the label ‘output hypothesis’, tended to be interpreted as a *product* rather than a *process*, even though she said the hypothesis was “about what learners did when pushed, what processes they engaged in” (p.473). Swain (2000) used the term ‘*collaborative dialogue*’ to emphasise the output process as both a cognitive and social activity where learners use language to mediate learning.

Collaborative dialogue then creates potential for the interlocutors’ utterances to

become an 'object' to be further explored (Swain, 2000, 2001, 2006; Swain, Brooks, & Tocalli-Beller, 2002). This is well captured in the words of Swain (2006):

Through speaking, thought is externalized. Externalized as an utterance, it becomes an object. As an object it can be scrutinized, questioned, reflected upon, disagreed with, changed, or disregarded. In order to collaborate, learners must speak to each other. Through their dialogue, they engage in making meaning, and debate the meaning made. To make their meaning as clear, coherent and precise as possible, learners will debate language form. (p.286)

This provides an alternative view on how form is brought into focus during the process of meaning making. Swain (2006; Swain & Deters, 2007) further introduced the term 'languaging' to refer to this process of meaning making as a dynamic process where language use mediates language learning.

Third, I look at the role of L1 use in mediating L2 learning. Research has shown that teachers often show unwillingness to use pair/group work for fear of students' use of L1 (e.g., Alley, 2005; Swain & Lapkin, 2001). Teachers also raise concerns about the usefulness of task-based learning in Asian EFL contexts. For example, they doubt whether students learn anything given their use of L1, among other factors (e.g., Carless, 2008; McDonough, 2004). Student L1 use has also been reported as one of the contextual constraints in teacher task implementation in Asian EFL classrooms (e.g., Butler, 2011).

Such are pedagogical concerns, yet research into pair/group work has demonstrated important mediating functions of L1 use in immersion, ESL and EFL settings. For example, Behan and Turnbull (1997, cited in Swain & Lapkin, 2000) investigated the use of L1 (English) by immersion Grade 7 students of French when they prepared for an oral presentation. The students worked in groups of four, obtaining information that each had on the lifestyles and environment of French natives. They were asked to use French L2 to do the task, but in two conditions: 'monitored' and 'non-monitored'. In the former, students were reminded to switch back to L2 when they fell back on L1 use. In the latter, they were not when they did so. Results indicate that the group who were not monitored had better oral presentations, and were able to transfer instances where

they used L1 to manage the task, and exchange information and search for L2 words to the presentation. This result, according to Swain and Lapkin (2000), shows a paradoxical and interesting role of L1. Behan and Turnbull concluded that L1 use functions to assist and promote language development as well as functioning as a cognitive tool in demanding tasks.

The question of how much L1 was used in different tasks were addressed by Swain and Lapkin (2000) in the context of Grade 8 students in French immersion classes in a Canadian school. Dyadic talk in preparation for two written tasks (a jigsaw and a dictogloss) was analysed. They found that students used L1 at 29% of the turns in the jigsaw task and 21% in the dictogloss task. Overall, of all the L1 turns produced, only 12% were off task talk. Notably these students used L1 mainly to move the task along, and do lexical searches. Swain and Lapkin concluded that their immersion teachers were “misinformed” when shying away from pair/groupwork (p.268) and that “to insist that no use be made of the L1 in carrying out tasks that are both linguistically and cognitively complex is to deny the use of an important cognitive tool” (pp.268-269).

Storch and Wigglesworth (2003) also found their ESL learners used L1 to a limited extent and for similar functions found in the above studies. Research in EFL contexts has also shown useful roles of L1 including (1) providing scaffolding, and establishing ‘inter-subjectivity’ (Antón & DiCamilla, 1998; Brooks & Donato, 1994; Villamil & De Guerrero, 1996); and (2) regulating and gaining control and raising awareness about their knowledge (Brooks, Donato, & McGlone, 1997). In a recent study in a college in Saudi Arabia, Storch and Aldosari (2010) found modest amounts of L1 use, at 7% for L1 words, and 16% for L1 turns when their EFL learners carried out three writing tasks: jigsaw, composition and text-editing in pairs. Consistent with the findings from previous research, students used L1 largely for managing the task, giving explanations on L2 vocabulary and conducting private speech, speech ‘directed to self’ (Centeno-Cortés & Jiménez, 2004). Research also shows that learners drew on L1 to carry out private speech as a language problem-solving strategy (e.g., Centeno-Cortés & Jiménez, 2004; Ohta, 2001).

The studies so far have shown a modest amount of L1 use. However, Guk and Kellogg (2007) found a high amount of L1 use by their Korean EFL primary school learners, at 46.93% of the total utterances produced in the context of five lessons which included groupwork subsequent to the teacher-led sessions. Alley (2005) also found high school students studying Spanish as an L2 in America used English L1 predominantly in groupwork, at 71%, though for different mediating functions in their project groupwork. This substantial L1 use also found support in Alegría de la Colina and García Mayo (2009), in a study on L1 use by undergraduate EFL low proficiency learners who carried out three collaborative tasks (jigsaw, text reconstruction and dictogloss), at 55-78% (calculated out of L1/L2 words) depending on the tasks. The amount of L1 use is clearly influenced by the task, learner proficiency and learning contexts. How much students use their mother tongue might also be influenced by their attitudes towards L2 use (Storch & Wigglesworth, 2003). However, research to date has not adequately addressed why students choose to draw on the native language (Lantolf, 2000). In the current thesis, student voices on why they use L1 during task talk were also documented and discussed.

Fourth, I look at the importance of learner agency in socio-cultural theory. This theory emphasises that it is the learners who take actions to realise the set goal (Donato & McCormic, 1994, p.455). Regarding focus on form, it is the individual learner who approaches their language problems, analysing and weighing their language solutions during the 'linguaging' process (Swain, 2006). In this way, how learners internalise language forms depends on their *agency* (Brooks & Swain, 2009), and how he or she is "afforded and constrained by her or his ZPD" (Lantolf & Thorne, 2007, p.266). In other words, it is not so much tasks that create environments for learning but the 'activity' the learner engages in that is important (Coughlan & Duff, 1994). However, Ellis (2003, 2012) argues that accepting the role of learner agency does not necessarily refute the role of tasks: tasks can have certain predictable influences (e.g., Bygate, 1999a; Newton, 2013; Newton & Kennedy, 1996; Skehan & Foster, 1997). Evidence for this can be seen through the effects of tasks on LREs even from studies that take a socio-cultural perspective (e.g., Alegría de la Colina & García Mayo, 2007; Storch, 2001a; Swain &

Lapkin, 2001). This is also the stance I will argue, as supported by the findings in this thesis.

In brief, a socio-cultural perspective on focus on form emphasises the important mediating role of language in L2 learning, the ZPD, L1 use and learner agency. The centrality of socio-cultural theory is summarised in the words of Swain (2005):

Socio-cultural theory ... puts language production in a “star role”, so to speak. Speaking (and writing) are conceived of as cognitive tools-tools that mediate internalization; and that externalize internal psychological activity, resocializing, and recognizing it for the individual; tools that construct and deconstruct knowledge; and tools that regulate are regulated by human agency. (p.480)

2.4.3 A cognitive approach

A cognitive approach offers an alternative perspective on how learners attend to language form during task performance and how this serves language learning. It is represented by two main theoretical models: the trade-off hypothesis (Skehan, 1998, 2009) and the cognition hypothesis (Robinson, 2001, 2007, 2011a, 2011b).

Because of the L2 user’s limited attentional capacity, Skehan (1996a, 1998) proposes the trade-off hypothesis which argues that three goals of language production, accuracy, complexity and fluency, compete for attention. In this view, L2 knowledge is represented in two systems: exemplar-based and rule-based. The former consists of lexical items or formulaic chunks of language which can be quickly accessed for use. The latter is composed of language rules or ‘abstract representations’ which need more control to be used, and which therefore place greater demands on the learner’s limited attentional capacity. It follows then that under the stress of communication, learners might necessarily opt for exemplar-based production, thus prioritizing fluency. In such circumstances, they have insufficient attentional resources to resort to rule-based systems to restructure their interlanguage (boost complexity) or to conform to target-language use (enhance accuracy). Fluency, complexity and accuracy therefore compete so that attending to one causes a trade-off in the others, especially between complexity and accuracy. Therefore, in order to achieve a ‘balance’ of these three goals, tasks

should be selected with reasonable difficulty by taking into account three factors: code complexity (e.g., linguistic task input), cognitive complexity (e.g., task topic, task familiarity, processing demands) and communicative stress (e.g., time pressure) (Skehan, 1998).

However, based on a parallel processing model, the cognition hypothesis (CH) (Robinson, 2001, 2005, 2007, 2011a, 2011b; Robinson & Gilabert, 2007) argues that both complexity and accuracy can be achieved concurrently without any cost to either via manipulation of task complexity. The CH distinguishes two dimensions of task complexity: resource-directing and resource-dispersing. The resource-directing dimension makes cognitive/conceptual demands but “*direct*” learners’ attention to relevant L2 features. For example, asking learners to do a task in a ‘there and then’ condition will direct their resources to the use of past tenses. The resource-dispersing dimension (e.g., removing planning time) *does not direct* learners’ attention to specific L2 forms, but “*disperses*” their attention over many L2 aspects. The CH predicts that increasing task complexity along the resource-directing dimension will push learners to extend their interlanguage to encode increasingly complex concepts, thus leading to greater accuracy and complexity in monologic tasks. This draws on the work of Givón (1985, 2009) arguing that complex conceptual meanings are expressed by complex linguistic forms. In essence, in the process of speech production, learners map their conceptualisation or meaning to language forms (Levitt, 1989; Slobin, 2003), and thus more complex meaning to be made will direct learners to seeking relevant L2 forms to encode (Robinson, 2011a, 2011b). This also shows a multiple resources view (Wickens, 2007) that native speakers produce language simultaneously accurately and fluently and with complexity. On these grounds, Robinson also claims that making tasks more complex along the resource-directing dimension will lead to more interaction and uptake of input made salient through focus on form techniques (e.g., recasts) in interactive tasks.

On the other hand, the CH predicts that increasing task complexity along the resource-dispersing dimension places greater performative demands on learners’ cognitive resources and thus robs attention from the linguistic demands of the task. While these increased performative demands reflect real life performance (and so can facilitate

fluency development), they are predicted to lead to a reduction in the accuracy and linguistic complexity of learner production and therefore to fewer opportunities for learners to expand their linguistic resources in the L2 through task performance. It is in this aspect that both the trade-off hypothesis and the CH converge. The CH also identifies two other groups of factors which affect task performance: task conditions (e.g., whether the task is one-way or two-way) and learner factors (e.g., ability and motivation).

Research within the cognitive tradition has sought to manipulate task characteristics and task conditions (e.g., pre-task planning, public performance as post-task activity) to channel learners' attention to the three goals of language production (e.g., Foster & Skehan, 1996; Skehan & Foster, 1997, 2001) (see Ellis, 2009b and Skehan, 2009 for recent reviews). Research has also examined the effects of task complexity on interaction (Gilabert, 2007; Gilabert, Barón, & Ilanes, 2009; Kim, 2009; Révész, 2009; Révész, Sachs, & Mackey, 2011; Robinson, 2001, 2007; Robinson & Gilabert, 2007). Empirical evidence has shown support for both Skehan's and Robinson's hypothesis, though in the main in favour of the former (Ellis, 2000, 2009b; Skehan, 2009). Furthermore, Skehan (2009) argues that an interaction between task characteristics and task conditions, rather than task complexity *per se*, as Robinson argues, would lead to simultaneously increased complexity and accuracy. Further reviews on studies that investigated pre-task planning, from a cognitive perspective, especially rehearsal and public performance, will be discussed in the next chapter.

2.4.4 Summary: Focus on form approaches in TBLT

In brief, focus on form that occurs during the course of communicative tasks can be achieved from different approaches including psycholinguistic, socio-cultural and cognitive. Ellis (2012) writes:

There is a tendency to view these theories as incommensurate and for researchers to stake out claims in favour of their own paradigm through criticizing the theoretical foundations and limitations of the other paradigm. This seems to me an unprofitable way to proceed. (p.34)

Ellis (2000, 2003, 2012) then asserts that though different in their theoretical assumptions, these approaches should not be seen as incompatible, but complementary. Swain and Deters (2007) also argue for a “broader” and “balanced” view on L2 learning in order to understand it (p.381). This ‘theoretical pluralism’ provides alternative paradigms to gain insights into the complex phenomenon of language teaching and learning (Ellis, 2008). And it is this stance that I have adopted throughout the thesis.

2.5 Classroom task implementation

Given the theoretical rationales of TBLT and tasks, it is important to see how tasks are manifested in classrooms. While an extensive body of laboratory or quasi-experimental research has documented the role of tasks in SLA, less research has been carried out in classrooms under normal operating conditions. In fact, the ways tasks have commonly been studied have not reflected the ways they have commonly been used (Samuda & Bygate, 2008).

However, research is increasingly addressing this gap (Eckerth & Siekmann, 2008; Edwards & Willis, 2004; Van den Branden, 2006c; Van den Branden, Van Gorp, & Verhelst, 2007). For example, Berben, Van den Branden and Van Gorp (2007), in Van den Branden et al.’s (2007) book, investigated how three teachers implemented the same single task, a radio news bulletin task, designed by professional experts in their classrooms. The classrooms comprised multilingual students with Dutch as the medium of instruction. The task was transformed in different ways by both teachers and students. In particular, the teachers deviated from the intentions of the task designers and transformed the task in ways that suited their beliefs and perceptions. They allowed students to add elements of creativity or fun in their task performance. Berben et al. concluded that

A task should not be perceived as fixed entity, but rather appears to behave as highly flexible and kneadable material that can take on different existential guises as it passes through the minds, mouths and hands of different persons making use of it. (p.56)

In a recent study that addresses the same research context, Van den Branden (2009a) reports four studies on teachers' and learners' re-constructions of tasks in order to respond to their goals, expectations and preferences. He argues that despite unpredictability associated with task interpretations, tasks offer space for teachers and learners to construct learning to achieve set goals. Van den Branden claims that TBLT does not cause more chaos than any other teaching methodology. Rather it accepts how complex and unpredictable task-based language learning is. He contends that it is within this complexity and unpredictability that the role of teacher as supporter, organiser, and social conversation maker, is called for to lead and help learners "move about the pedagogical spaces" (Samuda, 2007) that tasks provide.

In another study, Andon and Eckerth (2009) investigated the use of tasks by four teachers in an ESL setting in the UK, by means of classroom observations and interviews. They found the teachers used tasks in four main ways. First, they used tasks as a means for students to communicate and negotiate meaning using their own language resources. Second, they used tasks that reflect both real life target activities and tasks that trigger real life processes of communication (both situational and interactional authenticity). Third, they used tasks with a clear outcome. Fourth, they used focussed tasks as a way of focusing on form. Overall, the pedagogic principles in TBLT were reflected in the teachers' teaching, though with modifications. "All four teachers experiment with different elements of TBLT, reject some of them, embrace others, and combine all of them with other pedagogical elements." (p.305). However, Andon and Eckerth warned against generalizing the findings because they described their research context as 'privileged', in that the teachers had access to TBLT literature.

Samuda (2001) also examined how a teacher created 'a need to mean' through task design, targeting modality (e.g., may, might, must, can, could) which led students to make connections between form and meaning. The research was carried out in an intact ESL university classroom of nine students in North America. Through the teacher's skilful prompting, students developed more natural use of the targeted items, and thus expanded their repertoire beyond a limited stock of set terms such as 'probably' or 'maybe'. This study shows the important role the teacher played in scaffolding the meaning-form mapping process.

Another strand of research that looks at tasks in action involves evaluation of TBLT as curricular innovation in various contexts. East (2012) investigated teachers' (and school advisors') understandings and perceptions of tasks in a national school curriculum renewal for learning foreign languages in New Zealand. By means of semi-structured interviews with 19 experienced teachers (and 8 school advisors), he found that his teacher participants understood and used tasks in different ways, though in the main for meaningful communication. They saw the immediate school context as the real world situation for task design with both situational and interactional authenticity, as the teachers in Andon and Eckert's study. The teachers also revealed that they used more structured tasks for junior students and more open-ended with senior students. The study provides rich insights into TBLT including both positive reactions and challenges.

Other evaluation studies mainly involve TBLT as EFL curriculum innovations particularly in Asia, including Vietnam (Barnard & Nguyen, 2010; Carless, 2004, 2007, 2008; Deng & Carless, 2009; Le & Barnard, 2009; Nunan, 2003; Zhang, 2007). These studies typically show limited uptake of TBLT in the classroom. For example, Le and Barnard (2009), in a small-scale investigation of the implementation of TBLT as a curricular innovation by three teachers in a high school in a rural area in Northern Vietnam, found that the innovation was not translated in the classroom. Instead, the teaching was predominantly traditional teacher-led and textbook-reliant. Pressure to cover textbook tasks, teacher proficiency, insufficient resources, and non-task-based examinations were among the teachers' explanations for their classroom teaching. Studies in other Asian contexts also found analogous results. For instance, in a series of studies into TBLT in primary and secondary schools in Hong Kong, Carless (2003, 2004, 2007; Deng & Carless, 2009) found that the teachers reshaped the innovation in their own ways by preference for a PPP paradigm or task-supported language teaching (Ellis, 2003). As Carless (2004) points out, "teachers mould innovations to their own abilities, beliefs, and experiences; the immediate school context; and the wider sociocultural environment" (p.659). Zhang (2007), in a study of TBLT in a Chinese primary school context, also found the teachers did not use TBLT in their classrooms and reported inhibitive factors such as large class sizes, insufficient hours of learning and a lack of institutional support. Within the Japanese EFL context, Burrows (2008) argues that TBLT is not effective in such a country

because it is unrealistic and even unreasonable for a language programme to pose more demands on learners. Overall, TBLT in Asia is moving from “adoption to adaptation” (Butler, 2011, p.43) with the adaption being towards, at best, a weak version of TBLT or task-supported language teaching (see Adams & Newton, 2009; Butler, 2011; Littlewood, 2007 for recent reviews).

Despite such challenges, some positive reaction has been found. Hood, Elwood and Falout (2009), in a Japanese EFL context, found that students adopted positive attitudes towards TBLT, documented by their expressed willingness and comfort to communicate in English and understanding of the usefulness of this approach. Hood et al. then argue that the claim that TBLT is not appropriate in Japanese teaching contexts needs to be reconsidered by taking student voices into account. This positive reaction also found support in a study in a Thai EFL university context (McDonough & Chaikitmongkol, 2007). Although grammar instruction was identified by teachers and learners as necessary, both groups perceived TBLT to aid learners to develop their learning autonomy and address their academic needs. The study provides insights into the importance of the incorporation of learners’ needs and teachers in designing the syllabus.

What has been revealed in these studies into tasks in action is that teacher cognition - what teachers “know, believe, and think” (Borg, 2003, p.81) - shapes the choices they make in task implementation. Borg describes teachers as “active, thinking decision-makers who make instructional choices by drawing on complex, practically-oriented, personalised, and context-sensitive networks of knowledge, thoughts and beliefs” (p.81). Therefore, the personal practical theories that teachers hold have a powerful influence on their classroom practice (Borg, 2003, 2006, 2009; Pajares, 1992; Richards, 2008; Woods, 1996). These personal practical theories centrally tap into teachers’ idiosyncratic teaching *in situ* (Richards, 2008, p.167), which research needs to bring to the fore to inform language pedagogy in ways that help close the gap between theory and practice (Ellis, 1997, 2010b; Eraut, 1994). Furthermore, it is not only that the teachers’ practical theories shape their classroom actions but also that teacher actions and student actions are interdependent (Van den Branden, 2006b, 2009b). While this study is not a study of teacher cognition, this suggests the necessity to look at tasks in action from both teaching and learning perspectives.

2.6 Summary and link to the Phase 1 study

The chapter has discussed what TBLT entails, tasks, and psycholinguistic, socio-cultural and cognitive approaches to focus on form in TBLT. It also reviewed research on teacher classroom task implementation and discussed the role of teacher thinking. This line of research is valuable but has several limitations. First, studies that reported successful implementation of tasks in the classrooms (1) mostly documented the implementation level of *individual* tasks designed by experts and mostly in SL contexts (Berben et al., 2007; Samuda, 2001; Van den Branden, 2006c, 2009a) or (2) looked at tasks from a privileged context where teachers accessed TBLT literature (Andon & Eckerth, 2009) and from an institutional initiative (McDonough & Chaikitmongkol, 2007). Little research has analysed the recurrent modifications to the design of tasks made by teachers and why the teachers opt for such modifications within a set task-based curriculum. Such an analysis is likely to reveal important elements of teacher thinking that need to be taken into account in task-based curricula. Second, the majority of the EFL Asian evaluation studies investigated the uptake of TBLT as a national curricular innovation and found limited translation of TBLT into classrooms. Much less research has focussed on the actions and motivations of individual teachers implementing tasks, particularly teacher use of textbook tasks, teacher task design features and underlying rationales. Third, little research has linked teacher thinking underlying their task choice and implementation procedures to student learning, especially learning through oral tasks. As Borg (2009) noted, “a major issue that remains unaddressed is the relationship between teacher cognition and student learning” (Borg, 2009, p.169). Fourth, methodologically, no prior research investigated teacher use of tasks in *sequences* of textbook lessons, thus providing a limited picture of teacher tasks in action.

Phase 1 of the thesis addresses these limitations via an investigation of nine teachers using mandated textbook tasks in lesson sequences in EFL classrooms in a high school in Vietnam and teacher thinking underpinning their task choice and implementation procedures. In so doing it also responds to the pressing call for research with a classroom focus (Bygate, 2011; Bygate, Norris, & Van den Branden, 2009; Samuda & Bygate, 2008). It also explored how the students were engaged in classroom tasks (their perceptions and learning opportunities that occurred in tasks and task conditions

(rehearsal and performance) that the teachers used. In this way, the study further addressed the gap pointed out by Borg (2009) above: the link between teacher thinking and student learning.

The next chapter reviews studies on (1) rehearsal and public performance; and (2) language-related episodes (LREs) and language learning, which was the larger focus of Phase 2 of the thesis. The methodology for Phase 1 will be presented in Chapter 4.

Chapter 3 LITERATURE REVIEW 2: REHEARSAL, PUBLIC PERFORMANCE, LANGUAGE-RELATED EPISODES (LREs) AND L2 LEARNING

3.1 Introduction

In this chapter I begin by briefly reviewing an area related to the notion of public performance that arose from the teachers' task implementation in the Phase 1 study: Pre-task planning which includes strategic planning and rehearsal (Ellis, 2005). Then I discuss the notion of public performance and studies that investigate it. After that I review studies on LREs and L2 learning. Next I describe three main factors at issue in current LRE-focussed research, namely task type, proficiency and linguistic targets (lexical/grammatical). Finally I identify areas in need of further research which in turn are the larger focus of the Phase 2 study in this research.

3.2 Pre-task planning

Two types of pre-task planning have been identified: strategic planning and rehearsal (Ellis, 2005, 2009b; Skehan, 2009). They differ in that the former involves learners planning what to say and the language means to say it without rehearsing the entire task (Ellis, 2009b, p.474). Both share the same underlying assumption, from a cognitive perspective, that once learners have prepared what they want to say, they will have more attentional resources to attend to how to formulate and articulate the intended message (also see Bygate, 1996, 2001, 2005; Bygate & Samuda, 2005; Ellis, 2005). This assumption draws on the limited capacity model (Skehan, 1998) (see Chapter 2) and Levelt's (1989) model of speech production. The latter model identifies three main stages in speech production: conceptualisation, formulation and articulation. According to Levelt, conceptualisation entails preparing the message to be said, which then involves (1) deciding on the communicative goal; (2) carrying out macro planning by identifying speech acts needed; (3) retrieving relevant information; and (4) doing micro-planning to form a 'pre-verbal message'. This pre-verbal message is then sent to the formulator which then accesses and retrieves appropriate lexical items to be used and encoded in grammar. Formulation builds 'planned' or 'internal' speech, which is then

converted into overt speech in the final stage of articulation (see Bygate, 1996, 2001; Ellis, 2005, 2009b; Skehan, 2009 for interpretations of Levelt's model).

The following sections give a brief review of previous research into strategic planning and rehearsal.

3.2.1 Strategic planning

A large body of research has investigated strategic planning and found that it facilitates fluency and complexity of task performance while results on accuracy are inconclusive (Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996, 1999; Kawauchi, 2005; Mehnert, 1998; Mochizuki & Ortega, 2008; Ortega, 1999, 2005; Park, 2010; Sangarun, 2005; Skehan & Foster, 1997) (see Ellis, 2005, 2009b; Skehan, 2009 for recent reviews).

The inconsistent benefits of strategic planning on accuracy have led researchers to further look at three factors that may be at play: (1) Types of planning (teacher-led, group, individual; guided, unguided) (Foster & Skehan, 1999; Mochizuki & Ortega, 2008; (2) focus of planning (meaning/form) (e.g. Sangarun, 2005); and (3) task instructions (Park, 2010). Learner factors such as proficiency may also play a role. Ortega (1999) and Wigglesworth (1997) found that strategic planning enabled advanced learners to achieve greater accuracy in their task performance. However, Kawauchi (2005) found that given time to plan, low proficiency learners gained the most in terms of accuracy, while high proficiency learners benefited the most in terms of fluency and complexity, but advanced learners benefited the least. Kawauchi suggests that planning may have a ceiling effect at a certain level of proficiency. These differential advantages of strategic planning point to the need to find out what learners actually do when given time to plan, which has been rarely examined (Ellis, 2009b; Samuda & Bygate, 2008).

Ortega (2005) was one such rare study which investigated strategies that advanced and intermediate learners of Spanish used during pre-task planning by means of interview data. She found that intermediate learners were more concerned with retrieving lexical items, whereas advanced learners used a combination of strategies such as retrieval, rehearsal and monitoring. Despite this, some learners, both advanced and intermediate, reported what they planned did not transfer to their performance. It would seem then

that strategic planning does not always have a direct impact. Above all, striving for greater accuracy, complexity and fluency is a “learner-driven” or “learner-regulated” endeavour (Ortega, 2005). Learners’ orientations and attitudes towards the task and planning are very likely to have an influence (Ellis, 2009b).

On the differential effects of strategic planning, Bygate and Samuda (2005, p.39) claim that during strategic planning, speakers are more likely to attend to the content of what they say than to attend primarily to the linguistic resources required, thus focusing less on accuracy. By contrast, on-line planning in which learners are given time to plan during task performance has been found to push accuracy (Yuan & Ellis, 2003). This led Bygate and Samuda (2005) to argue:

... whereas strategic planning is likely to help speakers to prepare broader conceptual plans, and access mainly receptive language stores ..., its weakness is that whatever language plans might be accessed pre-task can turn out to be irrelevant or lost in the light of actual utterances. The opposite seems likely to be the case for on-line planning: this type of planning seems more finely tuned to the needs of specific upcoming utterances, but in this case, broader knowledge structures or language knowledge that is mainly reliant on controlled rather than already automatic processes may not be accessible. And this is simply because they have not been previously activated. (p.42)

Bygate and Samuda therefore proposed *task repetition*, as a form of planning, “integrative planning” or rehearsal (Ellis, 2005) to overcome the disadvantages of both strategic planning and on-line planning. The following section reviews studies on rehearsal.

3.2.2 Rehearsal

Several studies have examined task repetition, a form of rehearsal. For example, in a series of small-scale studies that focus on monologic task performance, Bygate (1996, 2001; Bygate & Samuda, 2005) found beneficial effects of task repetition. The 1996 study focussed on one student retelling a video-based story which was 2’50 long. The student watched a short video extract from a Tom and Jerry cartoon, (without any language input provided), recalled it and then retold it three days later. Bygate hypothesised that the repeated performance would be improved, since on the second

occasion the learner would be less concerned with conceptualisation and pay more attention to formulation and monitoring. Bygate analysed the lexical, grammatical and discourse features of the immediate performance (Time 1) and repeated performance (Time 2). He found (1) fewer errors; (2) more grammatically complex structures; (3) a wider range of vocabulary and (4) more self-correcting in Time 2. In brief, in Time 2 the student attended more to formulation and monitoring whereas in Time 1 he was more occupied with content work. In a later study that extended the 1996 one, Bygate (2001) examined the repeated performance 10 weeks later, of the same and new task of the same type by 46 international students at a British university. Two tasks were used: a narrative and an interview task. Bygate found that in the repeated task, students improved their fluency and complexity, though the difference in the general measure of accuracy was not significant. However, task repetition effects were not kept up in the new task, though of the same task type. In a follow-up re-analysis of the data from the 2001 study, Bygate and Samuda (2005) found a greater quality of performance in terms of propositional content and story explanations.

Like Bygate, Gass, Mackey, Alvarez-Torres, and Fernandez-Garcia (1999) in a study involving 103 English L1 learners of Spanish at an American university, also found greater fluency and complexity in a repeated task and no transfer to a different task, though with accuracy gain as regards the Spanish linguistic features- *ser* and *estar*. A lack of improvement in the new task, according to Gass et al., was probably because students might have been bored. In their words, “the novelty of the task may have ended and disinterest may have set in” (p.572). This contrasts with Bygate and Samuda’s (2005) claim that repetition does not necessarily reduce creativity, but might enhance it, because task repetition might offer a chance for students to use their “upper potential”(p.69).

The benefits of task repetition were also reported in Pinter (2005) with interactive tasks. Pinter asked 10 pairs of ten-year-old EFL primary school children in Hungary to do two information gap tasks: a ‘spot the difference’ task and a ‘follow the route on the map’ task three times. Each time students carried out similar tasks, not exactly the same tasks, with the same interlocutor and the time gaps between task repetitions were at least three days. Pinter found that students achieved greater fluency and accuracy over

the repeated performances. These benefits of task repetition were also verbalised in student reflections in a later analysis (Pinter, 2007). The students said they enjoyed the tasks and were aware of learning opportunities such as learning vocabulary from friends, and improving their repeated performances with the same partner.

Treating task repetition in a different way, Lynch and Maclean (2001) investigated task performance by 14 adult learners (oncologists and radiotherapists) carrying out a carousel task in normal classroom hours. The learners first prepared a poster in pairs, based on a medical article. No pre-task planning or rehearsal for the actual task performance was involved, as students spent time preparing for the poster only. Pair members then took turns to act as the 'host' and 'visitor'. The host had to respond to the visitor's questions. Each visitor presented a repetition of the task. Repetition in this study, according to the researchers, was immediate (every three minutes), with many "successive cycles" or "recycling" (p.143). Importantly, Lynch and Maclean found greater accuracy, complexity and fluency in the 'recycled' performances. However, while the higher proficiency learners were aware of the changes they made, the lower proficiency learners were not. Overall, like Bygate (1996, 2001), Lynch and Maclean argue that task repetition is a useful pedagogical option and needs 'selling' since repetition is not necessarily equated with boredom (cf. Gass et al., 1999).

In brief, the empirical findings on task repetition are encouraging. It improves learners' task performance and allows for tracing long-term learning (Bygate & Samuda, 2005). However, almost all the task repetition studies reviewed here were conducted in ESL contexts. In EFL contexts, students share a mother tongue and the relevance of oral communication tasks may be questioned given the traditional non-task examinations common in these contexts and a lack of an immediate need to communicate in the target language outside the classroom (e.g., McDonough, 2004; Pham, 2007). In such contexts, learners may need to be 'pushed' right there in the first performance (rehearsal). In this case, an impending public performance seems to be a useful pedagogical option to explore, and it is addressed next.

3.3 Public performance

Public performance has been proposed as a means to push learners to ‘perform’ their language at a higher level, and in so doing, to result in sustainable improvement in their L2 language proficiency (Skehan, 1996a, 1998; Skehan & Foster, 1997). It is also referred to as public report by Willis (1996) who sees it as one of the three phases of the task cycle (Task, Planning, Report) that occurs in the during-task stage (see Table 3.1).

TABLE 3.1: Willis’s (1996) task-based framework (p.52)

Pre-task
Introduction to topic and task
Task cycle
Task
Planning
<i>Public report</i>
Post-task
Analysis
Practice

Students first carry out the task in pairs or groups and the purpose here, according to Willis (1996), is to provide students with opportunities to use the target language with whatever resources they have to express their wanted meanings. The central focus is on building student confidence to produce extemporaneous discourse in private groups. In the planning stage,⁵ students plan and prepare for their public report, in which they briefly report, in front of the whole class, a certain aspect of the task “such as who won the game, how their group solved the problem, or two or three things they found out from each other” (Willis, 1996, p.55). Willis emphasises *reporting* the task outcome, not re-performing the task in front of the class (as used by the teachers in the current research). According to Willis (1996), a public report is

⁵ Planning here occurs after students have done the task, not pre-task planning.

the natural conclusion of the task cycle In itself it probably presents slightly less of a learning opportunity than the planning stage. But without the incentive of the report, the learning process of planning, drafting, and rehearsing would not happen. (p.58)

According to Skehan (1998), although Willis's (1996) framework provides teachers with a helpful guideline to push attention to form, it has four main limitations: (1) it lacks a theoretical foundation; (2) it does not link to how one's interlanguage can be developed; (3) as such it does not guide task and syllabus design (p.129); and (4) it lacks empirical evidence to support its use (also see Samuda & Bygate, 2008). From a researcher's perspective, Skehan (1996a) suggests public performance as a post-task option to pressure task performance as seen in Table 3.2.

TABLE 3.2: Skehan's (1996a) task-based framework (p.54)

Pre-task
Consciousness-raising
Planning
During-task
Task choice
Pressure manipulation
Post-task
e.g., <i>Public performance</i>
Own transcription

Skehan and Foster (1997, p.189, italics added) hypothesised that

Knowledge of a post-task, if the link between task performance and a subsequent public performance is clearly made in the learner's mind, can shift the fluency-accuracy balance very clearly towards the latter, since task performance itself, although unmonitored, will be seen as a *rehearsal* for the later performance where display and correctness of language assume greater importance.

Skehan and Foster studied task performance by 40 ESL students doing three tasks (a personal task, a narrative task and a decision-making task) in two conditions: one with 'foreknowledge' of a public performance and one without. They found that anticipation of an upcoming public performance promoted more accurate performance, but the

finding reached significance only in the decision-making task. They gave two explanations for this result. First, the data sample was small. Second, the learners may not make a strong link between the task and the post-task activity. Regarding the second explanation, better understanding may be gained by looking at learners' views on the public performance. Clearly public performance is worthy of further empirical investigation, as also pointed out by Skehan and Foster (1997), because our understandings of its effects are quite limited (p.207). No research since then has addressed this question. Although public performance has long been proposed, no research has documented whether and how teachers use it in classrooms, and how teachers and students perceive its benefits. Phase 1 of the thesis further provided insights into this neglected area.

Skehan and Foster (1997) only investigated the impact of anticipating a public performance on the accuracy, complexity, fluency of the actual task performance. To address this gap, Phase 2 of the current research investigates public performance from a process-product approach by examining language-related episodes (LREs) that learners produced during dyadic rehearsal and how language items targeted in these LREs were taken up in the public performance. The next section reviews studies on LREs and language learning.

3.4 LREs & L2 learning

This section provides an overview of the research findings of studies that investigate LREs in L2 learning. I will begin by distinguishing LREs from another related construct-focus-on-form episodes (FFE) and then review studies that have examined FFEs before focusing on LREs and L2 learning.

3.4.1 LREs & FFEs

LREs and FFEs have been used to refer to episodes in task-based interaction where student attention is drawn to language features in the context of communicative tasks by either the teacher or the students themselves (e.g., Ellis et al., 2001a, 2001b; Loewen, 2005; Swain & Lapkin, 1998, 2001). However, FFEs are typically used in focus on form studies that capture teacher-learner interaction (e.g., Ellis et al., 2001a, 2001b; Loewen,

2004, 2005; Nassaji, 2010). *Form* in FFEs can refer to grammar, vocabulary, pronunciation or discourse (Ellis et al., 2001b; Ellis, Basturkmen, & Loewen, 2002). Form refers “not just to form but also to the meaning(s) that a form realises, in other words, to form-meaning mappings” (Ellis et al., 2001b, p.415). FFEs are often confined to ‘interactionally accomplished’ FFEs, thus excluding student self-correcting (Ellis et al., 2001b, p.418).

LREs, in contrast, are typically documented in learner-learner interaction or collaborative dialogue (e.g., Kim & McDonough, 2008; Leeson, 2004; Swain, 1998; Williams, 1999, 2001) and are defined as

any part of a dialogue where students talk about language they are producing, question their language use, or other- or self-correct their language production (Swain & Lapkin, 1995). LREs thus entail discussion of meaning and form, but may emphasise one of these more than another. (Swain & Lapkin, 2001, p.104)

Despite these differences, FFEs and LREs share much common ground and some researchers (e.g., Zhao & Bitchener, 2007) have used them interchangeably. However, in this review I will keep FFEs and LREs separate in order to distinguish between form work that takes place in teacher-learner interaction (FFE) and in learner-learner interaction (LRE). A further reason for distinguishing LREs from FFEs is that FFEs focus mainly on correct resolutions (because the teacher provides feedback or answers to student queries) while LREs are typically coded as to the level of resolution (correct, incorrect and unsolved) (e.g., Swain, 1998).

3.4.2 Focus-on-form episodes (FFE) and L2 learning

Research has shown that both teacher and learner-initiated FFEs occur frequently in communicative ESL lessons, and that these FFEs target mainly vocabulary (Ellis et al., 2001a, 2001b; Loewen, 2004; Zhao & Bitchener, 2007). Research in EFL contexts (e.g., Farrokhi & Gholami, 2007) has also revealed that focus on form does occur, though with fewer student-initiated FFEs, at 16% in 24 hours of teacher-learner classroom interaction in an EFL Iranian university context. These studies also looked at the effectiveness of FFEs on immediate uptake (student incorporation of feedback in immediate utterances)

(Ellis et al., 2001a, 2001b). For instance, Ellis et al. (2001a) found a very high uptake rate of 76% of the FFEs in the two classes that they observed. The rate of successful uptake was also substantial in Zhao and Bitchener (2007), at around 53% for both teacher-learner interaction and learner-learner interaction. Farrokhi and Gholami (2007), in contrast, found a much lower uptake rate of 15.2%. Although these studies are valuable in showing the roles of both the teacher and students in bringing form into focus, they only report the effect of FFEs on immediate uptake. This is an important limitation since immediate uptake might be mere mechanical repetition rather than actual understanding, and a lack of uptake does not mean an absence of acquisition (e.g., Lightbown, 1998; Mackey & Philp, 1998; Panova & Lyster, 2002).

More recent studies have begun to address the link between FFEs and learning. For example, Loewen (2005) studied FFEs that occurred during 17 hours of communicative lessons in ESL schools in New Zealand. In the tailor-made post-tests, administered one day and one week after the occurrences of FFEs, Loewen found students scoring correctly 60% and 50% of the time respectively. In addition, the test scores were positively related to successful immediate uptake.

Similarly, Nassaji (2010) investigated FFEs in 54 hours of teacher-learner classroom interaction in seven ESL classes of three levels (upper beginner, intermediate, and advanced) in Canada. By means of tailor-made post-tests conducted one week after the interaction, Nassaji found that students scored correctly 72% of the time in the post-test and benefited more from the FFEs they initiated than those that the teacher did (72% > 46%). She also found that proficiency had little effect on the learning from student-initiated FFEs.

The benefits of focus on form were also reported in related studies with L1 Spanish learners of English (Alcón, 2007; Alcón & García Mayo, 2008). Alcón found that FFEs occurred often and mainly focussed on vocabulary. FFEs were linked to vocabulary learning measured by the learners' reported noticing in their learning journals and by written translation post-tests. Alcón (2009) further reported that of the reportedly noticed items, 40.2% were new lexical items, 52.7% were previously known items and only 6.7% were items where learner familiarity was not known. Additionally, in the three

post-tests (immediate, one week and three weeks after the FFEs), students used more new lexical items than previously known items, though the difference was not significant.

Overall, this body of research is largely limited to teacher-learner interaction and mainly to ESL contexts. It does not address the issue whether and how learners *themselves* attend to form in the course of communicative tasks in learner-learner interaction, especially in EFL contexts. As Williams (2001, p.304) argues, “if the effectiveness of FonF (focus on form) is ultimately determined by learner need, then it is essential to examine the episodes in which the learners themselves choose to focus on formal aspects of language.”

It is also important to see how students, not only by themselves, but also *among themselves*, bring form into focus during task-based interaction. Clearly when students interact with their peers, there is a different dynamic from that involved in interacting with the teacher (e.g., Long & Porter, 1985). A number of studies have investigated focus on form or LREs that arise in learner-learner interaction and L2 learning, and they are reviewed next.

3.4.3 LREs and L2 learning

In addition to studies that describe the occurrences of LREs in task-based collaborative dialogue (e.g., Fortune, 2005 ; Kim & McDonough, 2008; Malmqvist, 2005; Poole, 2005; Williams, 1999), considerable research has established an empirical association between LREs and L2 learning evidenced by means of different measures. Three main measures are identified: (1) tailor-made post-tests, (2) subsequent written task performance; and (3) subsequent oral performance which is broken into same task performance and spontaneous speech. Table 3.3 identifies research studies that used these three measures.

TABLE 3.3: LREs and language learning by various measures

Type of learning measures	Citations
(1)Tailor-made post-tests	Adams, 2007; Eckerth, 2008; Kim, 2008; La Pierre,1994, cited in Swain, 1998; McDonough, 2004; McDonough & Sunitham, 2009; Swain, 1998; Swain & Lapkin, 1998; Williams, 2001
(2)Subsequent written task performance	Brooks & Swain, 2009; Storch, 2002a, 2002b; Storch & Wigglesworth, 2007; Swain & Lapkin, 2002; Watanabe & Swain, 2007
(3)Subsequent oral task performance	
Performance of the same task	Bitchener, 2004; Donato, 1994; Truong & Storch, 2007
Spontaneous speech	Loewen, 2007; Williams, 2001

The next sections will review findings of the studies that have established an empirical relationship between LREs and L2 learning by means of (1), (2) and (3) above.

3.4.3.1 LREs and L2 learning measured by tailor-made post-tests

Tailor-made post-tests are a common measure of learning from LREs in research to date. Many studies using this form of measurement have focussed on learning items that learners *correctly* resolved during their task talk (Eckerth, 2008; Kim, 2008; McDonough & Sunitham, 2008; Williams, 2001). For example, Swain and Lapkin (1998) recorded collaborative dialogue by Grade 8 French immersion students while they were doing a jigsaw task in pairs. Each student had a different set of pictures, and they first created the story and then wrote it. Before students carried out the task, a five-minute lesson focusing on French reflexive verbs was introduced. Analysing the data from one dyad, Swain and Lapkin found that the dyad generated 23 LREs (15 grammatical and 8 lexical). Additionally, the number of LREs correlated positively with the post-test results and students tended to remember more language items that were the focus of LREs than items that were not.

Williams (2001) studied LREs produced by four pairs of ESL students of four proficiency (course) levels who carried out a range of classroom activities from meaning-focussed to more structured activities at an intensive English language programme in the United States. She also found an association between LREs and learning measured by a tailor-made post-test administered two weeks after the dyadic talk. The learners remembered 40-94% of the grammatical items and 50-94% of lexical items from correctly solved LREs, with higher proficiency learners scoring higher. Williams concluded that LREs between students were facilitative of learning.

In another study, McDonough and Sunitham (2009) examined LREs discussed by Thai EFL university learners doing a variety of collaborative computer self-accessed activities. The learners discussed more lexical LREs than grammatical LREs (76% > 24%) and resolved a large majority of their LREs (70% for lexical LREs and 84% for grammatical LREs). However, the retention rates in the post-tests were low, 48% for lexical LREs, significantly higher than 28% for grammatical LREs. McDonough and Sunitham gave two explanations for these low learning outcomes: (1) that students resorting to L1 extensively might not have provided a useful environment to enhance L2 learning; and (2) students might have thought the LREs discussed were not worth remembering because of an absence of follow-up activities and of their friends (not the teacher) providing the language solutions (p.249).

In a pre-test-post-test study, Kim (2008) investigated the effects of collaborative dialogue and individual work on vocabulary acquisition by intermediate Korean as a second language learners. Sixteen learners for each of the two groups, collaborative or individual, carried out a dictogloss task. The pre-test was composed of words taken out from the dictogloss passage. Based on a list of 20 words that students identified as being not familiar to them, in combination with additional questions which asked the learners to show whether they had known the word or not in the post-test. The study incorporated an immediate tailor-made post-test and a delayed one two weeks later. Kim found that students who worked collaboratively acquired (at least understood lexical meanings) at 70% and 74% of the lexical items discussed in the LREs in the immediate and delayed tailor-made post-tests respectively.

The findings of the above research are useful as they have shown the effectiveness of LREs and L2 learning. However, these studies focussed on only correctly solved lexical LREs. Little is known about how the language content of incorrectly solved LREs or unsolved LREs was retained in the post-tests.

Some studies have begun to address this. La Pierre (1994, cited in Swain, 1998) in a study of Grade 8 French immersion students carrying out a dictogloss task in pairs. LREs were first identified in pair talk. In the LRE-specific post-test administered one week after the LRE occurrences, students scored correctly 80% for items from 140 episodes in which they resolved their language problems correctly, and incorrectly 70% for items from 21 incorrectly solved LREs. These results support Swain's (1998) findings with the same task, in which students scored correctly 79% of the correctly solved LREs, and 29% of the incorrectly solved LREs. This indicates learning had occurred, even for the 'wrong thing' (Swain, 1998).

These results also find support in Adams (2007) in a study focussed exclusively on feedback episodes, generated by 25 students in an ESL American setting. Students performed three collaborative tasks that targeted past tenses, question-formation, and locatives. Using tailor-made post-tests, Adams found that (1) learners scored correctly on 59% of the post-test items; and (2) although the number of incorrect feedback episodes was small (20), more than half led to incorrect scoring in the post-test. Like the previous studies, Adam argues that 'mis-learning' does occur, though not predominantly.

Eckerth (2008) investigated collaborative dialogue in normal classrooms with a text reconstruction task (dictogloss) and a text repair task. The learners were lower and upper intermediate adult learners of German as a second language. Besides tracing learning gains of language points targeted by the tasks, Eckerth (2008) found that (1) a large majority of learning gains (78%) were related to LREs resulting from collaborative talk; (2) learners rarely changed a target-like item discussed in LREs into a non-target form in the delayed post-test; and (3) one third of incorrectly solved LREs led to correct scoring in the post-tests. These results led Eckerth to argue that incorrect learning via collaborative dialogue is rare. However, as Eckerth noted, due to the delayed post-tests,

students might have gained knowledge after the interaction, thus calling for cautious interpretations of the findings.

To summarise, the previous studies have shown that (1) collaborative work is conducive to learning as measured by tailor-made post-tests; and (2) whether learners resolve their language problem correctly or incorrectly, both lead to learning. Although tailor-made post-tests have the advantage of reflecting closely the content of LREs (Swain, 2005), they typically involve discrete language items and unlimited response time, thus presenting a mismatch between the nature of test items and LREs that occur in the context of communicative tasks (e.g., Adams, 2007; Loewen, 2007; Nassaji, 2010).

3.4.3.2 LREs and L2 learning measured by subsequent written performance

Research has also reported learning from LREs through subsequent written task performance. In Storch (2002a, 2002b), 33 ESL intermediate students from various L1 backgrounds at an Australian university carried out three collaborative writing tasks in pairs: a composition task, an editing task and a text reconstruction task. In the following week, they were told to do a similar writing task, but *individually*. The studies aimed to forge links between patterns of interaction and language learning. Storch identified four patterns of interaction: collaborative, expert/novice, dominant/dominant, and dominant/passive. 'Transfer of knowledge' was evidenced when in the individual written performance learners (1) used a new lexical item or a structure discussed in pair dialogue; (2) consolidated the existing knowledge of the vocabulary or structure resolved in pair interaction; or (3) used what they already knew in new contexts (Storch, 2002a, p.314). The 2002a study focussed on one task, the composition task, and found that the collaborative pair had seven instances of transfer of knowledge from collaborative talk to individual writings, while the dominant/dominant pair had none.

In the 2002b study, with a more detailed analysis with the three tasks, Storch found more instances indicating transfer of knowledge with the collaborative pairs (22 instances) and expert/novice pairs (15 instances) than with the other pairs (dominant/dominant and dominant/passive) (6 instances each). However, the dominant/dominant pairs had the most instances (8) of 'no transfer of knowledge' and

the dominant/passive pair had the most instances (7) of ‘missed opportunities’- instances where students made errors on their individual performance of the language items that they had little deliberated on in the preceding dyadic talk. Storch also found 10 instances where incorrect language resolutions in the pair talk were transferred to the individual written performance. Storch’s studies are valuable in showing the importance of the relationship between types of interaction and language development. Like Adams (2007), Swain (1998) and La Pierre (1994) above, this study has shown incorrect learning can happen during peer collaborative dialogue.

Swain and Lapkin (2002) also measured L2 learning by means of a subsequent individual writing task. They investigated LREs and learning that occurred when Grade 7 immersion students in Canada carried out a multiple-stage writing task. Their study design is summarised in Table 3.4.

TABLE 3.4: Swain & Lapkin’s (2002) study design

Stage 1 (Pre-test)	Original writing with visual prompts (jigsaw) in pairs, first orally and then in writing (30 minutes)
Stage 2	Comparing with the reformulated version (reformulation as feedback) (10 minutes)
Stage 3	Stimulated recalls where students talked about what they noticed between their original writing and the reformulation (40 minutes)
Stage 4 (Post-test)	Revising or rewriting the original version individually (15 minutes)

Swain and Lapkin (2002) found that in the original writing the pair discussed 47 LREs and 78% of these LREs were correctly solved. The results showed that 80% of the changes the pairs made were correct and of these correct changes, two-thirds corresponded to the exact reformulated items and one third were alternative expressions. When students accepted the reformulated items, most of them were correct and when they rejected the reformulations they were also correct 75% of the time. Swain and Lapkin argue that internalisation occurred due to the “talking it through” in many stages of collaborative writing. However, this study involved multiple stages and intensive time. It

is therefore unclear whether or not students can achieve the same learning outcome when they do collaborative oral tasks by themselves, as typically occurs in classrooms.

In a similar design study, Brooks and Swain (2009) asked students to work in dyads to explain and write a story based on a picture, 'the scene of the crime' in any way they wanted. Brooks and Swain found that students maintained, in their individual writing, from 90% to 98% of the correctly solved language items that were initially addressed through problem-solving. Furthermore, where students made changes to the formulated items, a large majority were correct: 64%-82%. Brooks and Swain concluded that learning resulted from the combined sources of 'expertise': peer, reformulator, and peer plus reformulator. Separating peers as a source of expertise, the study found that students were a reliable main source of knowledge for each other.

In brief, consistent with the results gleaned from the studies that measured learning via tailor-made post-tests, research has evidenced learning of LRE-related language items by means of a subsequent writing task performance. Because the nature of written and oral tasks is different (e.g., Kuiken & Vedder, 2012), one wonders how learning from LREs might also be manifested in subsequent *oral* communicative performance, and this is addressed next.

3.4.3.3 LREs and L2 learning measured by subsequent oral performance

Research that has provided evidence of learning via a subsequent oral performance falls into two groups: performance of the same task and spontaneous speech.

Regarding performance of the same task, three studies are worth mentioning. Bitchener (2004) investigated negotiation of meaning by 15 dyads of pre-intermediate ESL learners at a New Zealand university, doing an information gap and a decision-making task. One week and twelve weeks later, the students were asked to do the same task again, but with a different partner. Results showed that students used correctly around 70% and 62% of the correctly negotiated language items in the immediate and delayed repeated task performance respectively (and 76.8% in the tailor-made post-test, administered 12 weeks and 3 days after the first performance).

Donato (1994) studied collaborative dialogue by a group of French university learners who carried out one-hour pre-task planning to prepare for an oral presentation that occurred one week later. In the subsequent individual presentation, the learners were required to present to the class the conclusion of a scenario where a wife found out that her husband had bought a fur coat for another woman. Donato found that the learners used correctly, in their presentation, 24 out of 32 language items (75%) that they successfully resolved during the planning session. He concluded that language learning had occurred from dialogically mediated to independent performance. However, the delayed presentation might have encouraged students to invest extra individual efforts outside the collaborative planning to deliver it, thus weakening the claim that learning was due to the preceding social interaction (also see Eckerth, 2008). Furthermore, because this was a small-scale study (one group, doing one planning session) and the pre-task planning was intensive (one hour), its findings should be interpreted with care.

In a similar study, Truong and Storch (2007) examined the effects of collaborative pre-task planning on L2 learning. Five groups of four to five Vietnamese EFL university students were allowed 20 minutes to prepare for a subsequent *immediate* individual presentation on a given topic. The study found that (1) during the group planning, students were primarily concerned with content, with few LREs discussed (maximally 14LREs); and (2) very few LRE-specific items were taken up in the presentation (1-4 out of 14 LREs). Truong and Storch explained that the learners paid little attention to form for two reasons. First, the presentation in their study only provided a context for language practice, not for evaluation. Second, the pre-task planning was unguided. However, individual presentations may have encouraged individual rather than collaborative efforts, reducing co-responsibility. Thus a small number of LREs was discussed and uptake was limited. Subsequent public performance by the same dyad/group may bring about different results. Furthermore, learner proficiency and the task are likely to be additional contributing factors.

Only a few studies to date have used spontaneous speech as a measure of learning LRE-specific language items and found limited evidence of learning. For example, Williams (2001), mentioned earlier, found that the use of LRE-specific items in subsequent spontaneous interaction was low, 3-6 instances or 8-11% (though learners remembered

40-94% of the items in the tailor-made post-tests). Note that Williams recorded any occurrences of students happening to use the LRE-specific items randomly in subsequent classes where no specific task was used (Williams, 2012, personal communication via email 21-1-2012). She said this probably accounted for such a low occurrence because there was no particular work to elicit it. Similarly, Loewen (2007), in an extended analysis of Loewen (2005) reviewed earlier, found a low proportion of LRE-specific items being used in the subsequent spontaneous speech (19.8 % or 24/121 instances). Like Williams, Loewen (2007) explained that “... a lack of use of the targeted forms does not necessarily indicate an inability to use those forms; it may simply be that learners had no occasion to use them” (p.114).

Although it is difficult to elicit the target forms in non-obligatory contexts, the above findings may suggest that subsequent performance of the same task may better motivate internalisation than performance of unrelated tasks. McDonough and Sunitham’s (2009), for example, argue that the students in their study simply had “little incentive” to remember the language points they had discussed, because there were no post-task activities (p.248). Having said this, Truong and Storch (2007) above found low occurrences of LREs and limited evidence of learning in the subsequent performance. Other possible factors that may help explain these contradictory findings are tasks and learner proficiency. These and another factor, the linguistic focus of LREs, are discussed next, and each is addressed specifically in Phase 2 of the thesis.

3.5. Task types and LREs

In this section, I briefly discuss research studies on the effects of task type on negotiation of meaning, and on LREs, the focus of the current research.

3.5.1 Task types and negotiation of meaning

Certain tasks with particular characteristics can prompt learners to process and use language in ways that are useful to language acquisition (Ellis, 2000). As mentioned in the previous chapter, tasks such as closed (versus open) tasks, convergent (versus divergent) tasks, two way (versus one-way) information gap tasks and split (versus shared) information tasks have been argued to create more favourable conditions for

acquisition, because of the greater amounts of negotiation of meaning they induce (see Ellis, 2003, 2008; Mackey, 2012; Pica et al., 1993 for reviews). However, the superiority of these tasks has been challenged in different ways.

First, the value of tasks may vary, depending on learner factors (Mackey, 2012). Nunan (1991) argues that closed tasks might be more effective with lower proficiency learners. Similarly, Lambert and Engler (2007) suggest that closed tasks might be of limited benefit for advanced learners because they are more likely to constrain creativity. Interestingly, Julkunen (2001, cited in Mackey, 2012) found that closed tasks with a tight structure might suit anxious learners. All this seems to suggest a mediating role of learner variables.

Next, open tasks can offer potential, considering the fact that “learners need to align their resources in the expression of their own individuality” (Bygate, 1999a, p.187). In a study of negotiation of meaning in native speaker-non-native speaker interaction, Nakahama, Tyler, and Van Lier (2001) found that an open task, a ‘conversational activity’ created more opportunities for rich multi-level interaction than a closed information-gap task, though the latter induced more negotiation of meaning. Willis (2004) also suggests that open tasks may offer learners more opportunities to control the topic and handle their own talk. In open tasks learners are more likely to produce and sustain longer turns (Skehan, 1998). Indeed, Duff (1986) found learners produced longer turns and more complex language in debate tasks (divergent) than in problem-solving tasks (convergent), though the latter induced more negotiation of meaning (also see Nakahama et al., 2001).

The idea that ‘the more negotiation of meaning the better’ has also been challenged. For example, Foster (1998) found negotiation of meaning occurred infrequently in EFL classrooms. In a later study (Foster & Ohta, 2005), the results showed that learners attended to form in the course of doing communicative tasks in many other ways rather than negotiation of meaning to resolve comprehension problems. These included self-correcting, correcting others and scaffolding. Slimani-Rolls (2005) investigated negotiation of meaning in the context of an information gap task and a decision-making task. She found that quantitatively there was more negotiation of meaning in the former

task than the decision-making task, confirming previous studies. However, the individual data varied a great deal across task types. In addition, by investigating learners' idiosyncratic orientations and perceptions, it was shown that sometimes they did not engage in negotiation of meaning because they were able to fake comprehension and they wanted to avoid irritating their peers (see also Aston, 1986).

Finally, Newton (2013) argued that the nature of negotiation needs considering. In a study of four groups of ESL learners negotiating unfamiliar words in the contexts of performing two information gap tasks and two opinion gap tasks, he found more negotiation in the former than in the latter. However, students negotiated more word forms (e.g., spelling and pronunciation) in the former task type, but more word meaning in the latter. So the value of one task over another needs to be judged with caution.

In summary, previous research has shown that certain tasks induce greater quantities of meaning negotiation than others, as the case for closed over open tasks or convergent over divergent tasks (but see Newton, 2013 for qualitative changes). However, negotiation of meaning is only one of the many ways that learners attend to language form during the event of task execution. Open divergent tasks can also be conducive to learning in different ways. Clearly, more research is needed to investigate the potential of open tasks.

3.5.2 Task types and LREs

Recent research has shown task type influences not only the frequency of LREs that arise during learners' task talk, but also their linguistic foci (grammatical/lexical) and how they are resolved. For example, Storch (2001a) studied LREs that arose in pair talk when ESL students at an Australian tertiary institution carried out three text-based tasks, a text composition task, a text reconstruction task and an editing task. She found the editing task elicited the most LREs (490), followed by the text reconstruction task (410) and the last the composition task (209). However, the text reconstruction task elicited more grammatical LREs (92%) than the composition task (55%), and the editing task (73%), whereas the composition task induced more lexical LREs (39%) than the editing task and the text reconstruction task (25% and 5% respectively). García Mayor (2002) also found

that a text-construction task induced far more LREs than a dictogloss task (8 times) and that although students attended mainly to grammar in both tasks, they discussed more lexical items in the dictogloss task.

In a study of 12 pairs of ESL low proficiency students in a content-based programme of Spanish, Alegría de la Colina and García Mayo (2007) similarly found more grammatical than lexical LREs in all the three tasks used (jigsaw, text reconstruction, and dictogloss), though the jigsaw elicited more lexical LREs (28%) than the dictogloss (17.7%) and the text reconstruction task (14%). The text-construction task had the most LREs, followed by the jigsaw and then the dictogloss, but students were not most successful in resolving LREs in this task. The authors argue this was possibly because the LREs discussed in the text-construction task were beyond the learners' ability to solve.

Following a pre-test-post-test design, Swain and Lapkin (2001) investigated LREs discussed by two Grade 8 French immersion classes of similar levels. One carried out a jigsaw task and the other a dictogloss task. The pilot data with other classes were used to construct pre-test items. The post-test items included all the pre-test items and items that were created specific to LREs. Swain and Lapkin (2001) found more grammatical LREs than lexical LREs with both the dictogloss and the jigsaw. However, the two tasks did not differ significantly in the number of either lexical or grammatical LREs and in the post-test results. Swain and Lapkin explained that the limited number of matched pre-test-post-test items failed to capture learning pertaining to all the LREs. However, regarding the range of LREs, time completed and learner written performance, the dictogloss task was found to be more constraining than the jigsaw task. This was, according to the authors, due to the fact that the dictogloss task provided a language text while the jigsaw task did not. It could be hypothesised in the same way that more open-ended tasks such as debate tasks (divergent tasks) would be less constraining than problem-solving tasks (convergent tasks). These two task types were used in Phase 2 of the thesis.

Studies on the impact of task types so far have mainly focussed on text-based writing tasks such as dictogloss, text composition, text editing tasks, and jigsaw tasks. Very few studies have investigated the effects of different *oral* communication tasks on LREs.

Gilabert et al. (2009) studied the effects of task complexity on interactional processes such as negotiation of meaning, LREs, and recasts in the context of EFL students carrying out three tasks: a narrative reconstruction task, an instruction-giving map task, and a decision-making task. Gilabert et al. found no significant differences between these three task types in the frequency of LREs (and recasts and repairs), although the map task had the most instances of negotiation of meaning (clarification request, comprehension checks).

However, Gass, Mackey, and Ross-Feldman (2005) found an effect of tasks on LREs that arose in task-based interaction by 74 English L1 university learners of Spanish who carried out a picture difference task, a map task and a decision-making task in both laboratory and classroom contexts. The picture difference task involved students finding 10 differences between the pictures each had. In the map task, each dyad member had a different list of street properties and they had to locate the streets and draw a route. The decision-making task required students to agree on the ranking of a list of descriptions of Spanish universities in Spanish. Gass et al. found significantly more LREs with the picture difference and the map task than the decision-making task.

In brief, previous research has demonstrated that task types (written or oral) can influence the frequency with which students verbalise their language problems and the language focus of their problems. However, the focus of this research has been more on text-based or writing tasks than oral communication tasks. Few studies have addressed the impact of tasks on how learners resolve LREs (but see Alegría de la Colina & García Mayo, 2007), and no research has investigated the effects of oral task types on learning associated with LREs (but see Swain & Lapkin, 2001 for writing tasks).

Phase 2 of the present research investigated the impact of convergent and divergent tasks on LRE-related learning to address the teachers' choice of tasks that emerged from Phase 1 (see Chapter 5 and Chapter 7). Furthermore, since Duff (1986) investigated the effects of convergent tasks and divergent tasks on negotiation of meaning, little focus on form research has further explored these two task types, compared with a large body of research that has repeatedly used information gap referential tasks (as also noted by Skehan, 2003) (cf. Nakahama et al., 2001).

Methodologically, none of the LRE studies reviewed here controlled time for task completion, while time has been found to correlate with the frequency of LREs (Storch, 2001a; Swain & Lapkin, 1998). In addition, task topic is a potential variable, but it was not controlled in LRE studies on oral communication tasks (Gass et al., 2005; Gilabert et al., 2009)(cf. Newton, 2013), though it was in collaborative writing tasks (e.g., Alegría de la Colina & García Mayo, 2007; Swain & Lapkin, 2001). The Phase 2 study kept both time and task topic constant to minimise confounding effects.

3.6 Proficiency and LREs

In pair and group work where students carry out a given task, how to pair or group students of differing abilities to provide beneficial learning opportunities is a central pedagogical concern. Learner proficiency has been reported as one of the factors causing difficulty in classroom task implementation, particularly in EFL contexts (e.g., Butler, 2011; Littlewood, 2007; Pham, 2007). However, research into this proficiency variable in task-based interaction is still limited to date (see Philp & Tognini, 2009).

Some earlier studies have examined the effect of proficiency on negotiation and modified output (Iwashita, 2001; Yule & Macdonald, 1990). For example, Iwashita (2001) investigated the effects of proficiency on opportunities for modified output (interactional moves such as clarification requests) and amounts of modified output. Twenty-four students of Japanese as a foreign language at an Australian university carried out two tasks: one two-way information task and two one-way information tasks (describe-draw) in a 'meeting room'. The students were paired into high-high (HH), high-low (HL), and low-low (LL), four dyads each. Iwashita found that the HL dyads produced the most interactional moves, but not the highest amount of modified output. However, the three groups did not differ significantly in both measures. Iwashita therefore concluded that teachers may not need to worry too much if there are differing proficiency levels in the same L2 class.

However, Yule and Macdonald (1990) examined whether giving dyad members different roles would affect their interaction. They found that when a lower proficiency learner was assigned with a dominant role (e.g., sending information), more negotiation of

meaning ensued and they tended to successfully negotiate linguistic problems. On the contrary, when a higher learner undertook a dominant role, the lower one was passive in interaction, because the former often did not pay attention to what the latter said.

Recent research into collaborative dialogue in which learners discuss and resolve their language problems (Swain, 2000) has also begun to look at the effect of proficiency. Kowal and Swain (1997) found that in a mixed proficiency dyad, the less proficient learner was often dominated and ignored by the more proficient learner. Nonetheless, Storch (2001b) showed a quite contrasting result. Of the dyad participants in a text composition task, the heterogeneous proficiency dyad was more collaborative, thus having more LREs and more evidence of learning than dyads of homogeneous proficiency. Storch concluded that the nature of interaction may be more influential than a discrepancy in proficiency. This finding, to an extent, echoes that of Watanabe and Swain (2007) below.

In a recent laboratory-based study with a multiple-stage writing task (also see 3.4.3.2), Watanabe and Swain (2007) investigated the occurrences of LREs, patterns of interaction and learning outcomes measured by individual writing as a post-test. The data were collected from four “core” adult Japanese learners of ESL at a Canadian university, interacting with interlocutors of higher (core-high) and lower proficiency (core-low) while writing an essay on a given topic together. The results showed higher occurrences of LREs when students interacted with higher proficiency interlocutors than with lower ones and a higher overall learning gain for the core-high pairs (63%) than the core-low pairs (50%). However, the core students appeared to gain more when interacting with lower (64%) than with higher (58%). Watanabe and Swain argued that the nature of interaction is important. When pairs were found to be collaborative, they produced more LREs and scored higher on the written performance as post-test, irrespective of proficiency. The impact of the nature of interaction rather than proficiency alone is also supported by Dobao’s (2012) findings. In a qualitative analysis of dyad talk, Dobao found that an intermediate dyad had more LREs than an advanced dyad because they were more collaborative in their interaction.

Other research has shown a clear impact of proficiency on not only how often LREs occur, but also how they are resolved and associated learning. For example, Williams (1999, 2001) found that although learners focussed primarily on lexis (80%), higher proficiency learners discussed more LREs and achieved higher scores in the LRE-specific post-test. That higher proficiency seems to benefit learners also finds support in Kim and McDonough (2008), who investigated the task-based interaction by 24 university learners of Korean as a second language. Eight intermediate learners interacted with eight other intermediate and with eight other advanced learners. They carried out two dictogloss tasks of the same genre (biographies) and of the same lengths. The findings showed that (1) overall students discussed more grammatical LREs than lexical LREs and they were able to resolve correctly a majority of the LREs; (2) working with advanced interlocutors, students produced more LREs (both grammatical and lexical) than working with intermediate counterparts, though the difference was significant only for lexical LREs; and (3) students correctly solved significantly more LREs when working with an advanced learner (70%) than interacting with an intermediate partner (58%).

In another study that also used a dictogloss task, Leeson (2004) examined the effects of proficiency on the frequency, types and outcomes of LREs that occurred in collaborative dyadic dialogue. Twenty one dyads of adult Spanish learners in a content-based classroom carried out a dictogloss task. He divided dyads into three pairings: high-high dyads (HH) (8 dyads), high-low dyads (HL) (9 dyads), and low-low dyads (LL) (4 dyads) in regular classroom hours. Leeson found more LREs with increased overall dyad proficiency. Proficiency also affected the linguistic foci and outcome of LREs. The HH dyads had significantly more grammatical (67.11%) than lexical LREs (32.89%) while the HL attended to grammar and lexis almost equally, 54% and 46% respectively. The LL dyads focussed more on vocabulary (58.33%) than grammar in their LREs (41.67%). Leeson explained that because more proficient learners did not have as much difficulty comprehending the dictogloss passage, they had more attentional resources available to focus on grammar. Although all the three pairings resolved a majority of their LREs, the HH pairs correctly solved more LREs than the other groups (HL, LL), while the LL pairs left the most LREs unresolved.

In summary, previous studies have shown variable impacts of proficiency on how often learners discuss LREs, how they resolve them and subsequent learning outcomes. These studies have mainly used collaborative writing tasks (Kim & McDonough, 2008; Leeser, 2004; Storch, 2001b; Watanabe & Swain, 2007) or a wide range of activities (from tasks to exercises) (Williams, 1999, 2001). Furthermore, they have largely focussed on intensive ESL classes (Watanabe & Swain, 2007; Williams, 1999; 2001) and content-based Spanish programmes (Leeser, 2004) and in Korean L2 classrooms (Kim & McDonough, 2008). None have investigated how proficiency influences task talk in EFL contexts. Furthermore, the very few studies that have examined the effect of proficiency on LRE-related learning used tailor-made post-tests (Williams, 2001) and written performance (Storch, 2002a, 2002b; Watanabe & Swain, 2007). No research has explored how proficiency may mediate learning in subsequent public task performance. Not much has been known on how proficiency affects learning of different language aspects such as vocabulary and grammar. These themes are also addressed in Phase 2 of the thesis.

3.7 The linguistic focus of LREs and L2 learning

Some researchers argue that differing language areas such as lexis and grammar may require differing levels of processing and attention (e.g., DeKeyser, 1998, 2003; Doughty, 2003; Schmidt, 2001; VanPatten, 2004). Gass and colleagues (Gass & Alvarez Torres, 2005; Gass, Svetits, & Lemelin, 2003; Jeon, 2007) have demonstrated that linguistic focus may have an impact on L2 instruction. They assert that learners may need less ‘focussed attention’ to learn vocabulary than grammar. This argument is also in line with findings that show learners are more accurate in perceiving the intention of feedback on errors that involve lexical than grammatical items (e.g., Gass & Lewis, 2007; Kim & Han, 2007; Mackey, Gass, & McDonough, 2000). Research has also looked at the linguistic focus (e.g., lexical/grammatical) of LREs that learners discuss during task talk and how they resolve them (Kim & McDonough, 2008; Leeser, 2004; Poole, 2005; Swain, 1998; Williams, 1999). Very few studies have measured learning of these language aspects. McDonough and Sunitham (2009), reviewed earlier, showed learners retained more lexical (48%) than grammatical items (28%) in a tailor-made post-test. However, Williams (2001) found that learners, irrespective of proficiency, had almost equal

proportions of lexical and grammatical items scored correctly in the post-tests, although higher proficiency dyads scored higher in the overall test scores (see 3.4.3.1). Further research is obviously needed to investigate the effectiveness of lexical and grammatical LREs on L2 learning. Phase 2 of the current research further examined whether lexical and grammatical items were more likely to lead to learning in a rehearsal-public performance condition and how proficiency mediated this learning.

3.8 Summary and link to the Phase 2 study

This chapter has presented an overview of research findings on (1) pre-task planning (strategic planning and rehearsal) and public task performance; (2) FFEs/LREs and L2 learning by means of different measures; and (3) task types, proficiency and linguistic focus (lexical/grammatical) as mediating factors in LREs. The chapter has, along the way, identified the gaps that further motivated Phase 2 of the thesis. These gaps are summarised below.

First, research into public performance is rare, and yet it may be a common acting in many classrooms.

Second, no research to date has explored how task type and proficiency influence L2 learning in a rehearsal-public performance condition.

Third, very few studies have provided learning of language areas such as vocabulary and grammar. None have explored lexical and grammatical learning through oral communication tasks in a rehearsal-(public) performance condition

Fourth, the methodology of the majority of research that investigated the effects of task types on LREs did not control for time or task topic for task completion, thus making the results subject to confounding task effects.

Finally, research has largely focussed on measuring learning associated with LREs by means of tailor-made post-tests, though acknowledging a mismatch between the nature of these tests and that of LREs (Adams, 2007; Loewen, 2005, 2007). Many studies have also measured learning of LRE-specific language items through subsequent written performance, but very few studies have used oral task performance as a measure of

learning. In the current study, public performance, aside being used by the teachers as a stimulus for target language use, is a measure of learning that contributes to this gap.

Phase 2 of the thesis addresses the limitations of previous research by examining the effects of task type (convergent/divergent) and proficiency on the take-up, in dyadic public task performance, of language items that had already been attended to in LREs in task rehearsal. The two tasks were on the same topic and were carried out within the same amounts of (rehearsal) time for all proficiency groups (see Chapter 7). The study also looked at how the linguistic focus of LREs mediated learning. It further documented learners' task perceptions and the strategies they said they employed during rehearsal for the performance.

The next chapter presents the methodology for Phase 1, the results of which will be reported and discussed in the next two chapters before the methodology for Phase 2 is described in Chapter 7.

Chapter 4 THE PHASE 1 STUDY: METHODOLOGY

4.1 Introduction

In this chapter I will describe the methodology for the Phase 1 study. I will first introduce the research site, and the teacher and student participants. I will then present the data collection methods and data collection procedures. Next I will describe how the data were analysed. Finally, I will discuss how the issues of validity and reliability of the study were addressed and conclude with a summary.

Phase 1 examined how teachers used and implemented prescribed oral textbook tasks in their classrooms within the context of a new task-based English curriculum for Vietnamese high school students and the rationales behind their practices. It also explored how students engaged in the classroom tasks and their perceptions of learning opportunities through tasks. It addresses the following research questions (RQ).

Teachers using and implementing textbook tasks:

- RQ1. How closely did the teachers follow the textbook tasks?
- RQ2. In what ways did the design features of the teachers' tasks diverge from the textbook tasks, and why did they diverge from them?
- RQ3. What task implementation procedures did the teachers use, and why did they use them?

Students engaging in tasks:

- RQ1. To what extent did the Vietnamese high school students attend to form while rehearsing for the performance of communicative tasks? If so, how?
- RQ2. To what extent and for what purpose did the students use L1 in task rehearsal?
- RQ3. To what extent did the students use items accurately in performance that had been subject to LREs in rehearsal?
- RQ4. How did the students perceive communicative tasks, task rehearsal and performance?

4.2 The research site

The study took place at a high school in Vietnam, one of the elite long-standing high schools in the country where the researcher had taught for 13 years before taking PhD study leave. In Vietnam, high school students are at three grades (10, 11 and 12), and aged from 15 to 18. At the time of data collection, students attending this high school fell into two main strands: majors and generals.⁶ Majors include students who specialise in different subjects including mathematics, physics, chemistry, biology, literature, English, French, history and geography. These majors are grouped into classes according to their chosen subjects. General classes are composed of non-majors who do not specialise in any particular discipline. In the case of English, the school has classes for English majors and classes for non-English majors. The latter follow the national mandatory textbooks. This school is the only high school in the area that has majors and that requires an entrance examination for prospective students. The school is well-known for its students' academic achievements. Every year, many students from this school win national (and international) prizes in the examinations for gifted students. The school ranks high in terms of the rates of students successfully passing graduation and university entrance examinations in Vietnam. Many students from the school have also earned scholarships to study overseas. The teachers teaching at this high school typically have a high distinction undergraduate degree or are experienced teachers who have won teaching awards in teaching competitions organised by the local Department of Education and Training. The teachers during their teaching also pursue higher education, and many teachers now have a Masters' degree. In Vietnam, it is a mark of prestige for those who study and teach at this high school.

4.3 The teacher and student participants

Teachers and classes who used the new task-based textbooks were recruited for the study. There are two main streams of English textbooks for high school students in Vietnam, namely the General-English (*Tiếng Anh cơ bản*) and the intensive-English (*Tiếng Anh nâng cao*). My research solely focuses on the General-English textbooks and the

⁶ From 2012, the school has only recruited majors.

classes that used this textbook for two reasons. First, these textbooks are used by a majority of students in the high school and in wider Vietnam. Second, the textbooks espouse the communicative approach which points to *learner-centredness* and *task-based methodology* as the central focus (Hoang et al., 2006, 2007). This purposeful sampling (McMillan & Schumacher, 1993) was elected to seek knowledge and information about the phenomenon being investigated, that is, how teachers used and implemented textbook tasks and how learners engaged in classroom tasks.

At the time of data collection, the first semester of the academic year-2010, those English teachers who were in charge of English majors,⁷ were not teaching any other classes, leaving other teachers to be exclusively responsible for general classes and classes of other majors. Nine of these other teachers volunteered to participate in this Phase 1 study. Three of them taught Grade 10 classes, three Grade 11 classes, and three Grade 12 classes. These teachers taught classes that followed the task-based textbooks. Nine of the intact classes taught by these nine teachers also agreed to take part in this research. Table 4.1 provides a summary of the demographic information on both teacher and student participants.

⁷ The teachers who taught English majors could use additional materials of their own choice to prepare students for examinations intended for these classes.

TABLE 4.1: The teacher and class participants

Teacher					Intact classes			
ID	Years of experience	Qualifications	Gender	Age	Grade	Class type	Class size	Gender
110A	11	BA, enrolled in an MA in Applied Linguistics	F	34	10	Chemistry majors	31	Mixed(16F/15M)
210B	17	BA	F	43	10	Literature majors	32	32F
310C	Two months	BA	F	22	10	History and Geography majors	28	Mixed (14F/14M)
411D	3	BA	F	25	11	General class	46	Mixed(32F/14M)
511E	Two months	MA in TESOL	F	24	11	Maths majors	28	Mixed (9F/19M)
611F	11	BA	F	42	11	Information Technology majors	29	Mixed (8F/11M)
712G	15	BA	M	38	12	General class	40	Mixed(19F/21M)
812H	17	BA	F	39	12	Physics majors	27	Mixed(5F/22M)
912I	23	BA	F	47	12	Information Technology majors	26	Mixed(7F/19M)

Note. ID= Identification; F = female, M = male

As seen from Table 4.1, eight out of the nine teachers were female, which reflects the popularity of teaching English as a career for women in Vietnam (Ho, 2011; Nguyen, 2003). The majority of teachers, 6/9, were between 34 and 47 years of age and had been teaching EFL between 11 and 23 years. The three youngest teachers (310C, 411D, 511E) had from 2 months to 3 years of teaching experience. All teachers had a minimum

of a BA in English. The majority of the teachers (except 310C and 511E), attended training workshops in ELT techniques several times in their teaching, organised by VTTN (Vietnam Teachers & Trainers Network) in collaboration with the British Council in Vietnam. These teachers had also been trained in how to use the new series of textbooks in workshops held by the Ministry of Education and Training or the local Department of Training and Education.

Of the nine classes in the study, seven were majors of different disciplines including Chemistry, Literature, History and Geography, Maths, Information Technology, and Physics. Two classes were of the general type, that is, students were not specializing in any subject. Although all the classes differed in their majors, they were all using the same task-based textbook intended for their grade. All the classes had three 45-minute periods of English per week, which is usually split into two sessions: one single session of 45 minutes and one double session of 90 minutes. The students in each class were typically of a similar age, usually 16 in Grade 10, 17 in Grade 11, and 18 in Grade 12. The average class size ranged from 26-46 students, with general classes usually being more crowded. The great majority of the classes were mixed-gender.

In Vietnam English is officially a compulsory subject when students enter Grade 6, so by the time of data collection, Grade 10, Grade 11, and Grade 12 students had been officially learning English for 4 years, 5 years and 6 years respectively.

4.4 Data collection methods

Phase 1 is “descriptive research” which aims to provide “*qualitative and quantitative* accounts of classroom processes, the factors that shape these and their implications for language learning” (Ellis, 2012, pp.41-42, italics added). I chose the term descriptive research to name the Phase 1 study because it can include both qualitative and quantitative analysis, thus avoiding what Bryman (2008) describes as the problematic divide between quantitative and qualitative. Quantification of qualitative data brings about “a meeting ground” for the qualitative-quantitative traditions (Bryman, 2008, p. 598), which is also my approach to data analysis throughout. A ‘descriptive’ approach

does not mean that the research does not also explain the phenomenon under investigation (Ellis, 2012).

According to Ellis (2012), descriptive research (1) presents rich stories from a specific teaching and learning setting from the insider's perspective; (2) involves no interventions, but investigates phenomena as they are; (3) "often involves a research-then-theory approach" where there are no pre-conceived units; (4) shows that the results of the study are "trustworthy" and portrays the perspectives of different participants involved (p.42).

This Phase 1 study aimed to examine how the teachers used and implemented the textbook tasks in their classroom. The study then looked at how the students made use of the learning opportunities made available through the classroom tasks. In order to obtain the answers to the questions being sought, I exercised the rigour of "a fit between question, data and method" (Richards, 2009, p.17) in research design from developing research questions to data sources and data collection methods. This fit was realised through multiple data sources including classroom observation which was video and audio recorded, field notes, classroom materials, teacher interviews, and student interviews.

4.4.1 Classroom observation

In this study, classroom observations were carried out in classrooms in naturally occurring lessons. As Nunan (1992) points out, "the context in which behaviour occurs has a significant influence on that behaviour. It follows that if we want to find out about behaviour, we need to investigate it in the natural contexts in which it occurs..." (p.53). Furthermore, classroom observations allow 'live' accounts from 'live' settings (Cohen, Manion, & Morrison, 2007). The observational data also laid ground for behaviour to be explained. In Punch's (2006, p.34) words, "a good first step in explaining why something happens is to describe exactly what happens." In the current study, the descriptions of actual practices were obtained through classroom observations in normally scheduled lessons supplemented by video and audio recordings, field notes and classroom materials. These supplementary sources of data acted both as a triangulation and a

further tool (for interviews) to understand the teachers' practice and student perceptions.

4.4.1.1 Multiple case studies

This Phase 1 study adopted a multiple case study approach (Stake, 2005), involving classroom observations of nine teachers, three from each grade level (10, 11, 12), teaching the same textbook units. Stake argues that by illustrating "how a phenomenon occurs in the circumstances of several exemplars", this approach "can provide valued and trustworthy knowledge" (pp.458-459).

Each of the nine teachers and their classes were observed across the five lessons that make up a textbook unit, namely Reading, Speaking, Listening, Writing and Language Focus. The main data collection included the Grade 10 teachers teaching Unit 6, *An Excursion*, Grade 11 teachers teaching Unit 7, *World Population* and Grade 12 teachers Unit 6, *Future Jobs* (See Table 4.2) (see Appendix 3 for a copy of the textbook units).

TABLE 4.2: Classroom observation scheme

Teacher		Textbook unit	Lesson
Grade 10	110A	Unit 6: An Excursion	1 Reading
	210B		2 Speaking
	310C		3 Listening
			4 Writing
			5 Language focus
Grade 11	411D	Unit 7: World Population	1 Reading
	511E		2 Speaking
	611F		3 Listening
			4 Writing
			5 Language focus
Grade 12	712G	Unit 6: Future Jobs	1 Reading
	812H		2 Speaking
	912I		3 Listening
			4 Writing
			5 Language focus

In total, data were collected from 45 lessons. Although the focus of the research was on the oral tasks, the complete unit was observed to obtain a fuller picture of task-based teaching in these classes, as Samuda and Bygate (2008) argue

The interrelationships between a task, its position in a teaching sequence, the pedagogic role it plays within that sequence, and the purpose motivating its use are complex, and ... from a pedagogic perspective it is necessary to focus on understanding tasks in light of those relationships. (p.218)

During classroom observations, I took the role of a non-participant observer. I informed the teachers of my general research area, that is, how task-based language teaching was used in the classroom. However, to avoid the danger that the teachers would teach towards the data, the precise focus (oral tasks) of the research was not specified. The observation procedures were the same as described in the piloting section (see 4.5.2).

4.4.1.2 Video recordings

The lessons were video recorded because “film preserves activity and change in its original form” (Marshall & Rossman, 2006, p.121), thus allowing for repeated viewing of the lessons, and discovery and re-discovery of the phenomenon being researched. The videos allowed me to move back and forth between the recorded lessons to check the emergence of themes in the teachers’ use of textbook tasks and their task implementation against other data sources (e.g., teacher interviews, student interviews, student task talk).

The video recordings of the observed lessons were also later used in stimulated recall sessions with the teachers and students. The recordings were important because it was not always possible to conduct these sessions right after the observations as the teachers often had consecutive classroom hours with different classes on their teaching days, and the students also had consecutive classes in different disciplines. Also, since teachers (and students) might unavoidably re-construct their descriptions if reliant exclusively on memory (Breen, Hird, Milton, Oliver, & Thwait, 2001), video recorded lessons provided a reliable and practical context to probe the teachers’ rationales for their practice. In addition, in order to examine the teachers’ thinking which “is very

much concerned with teachers' *personal* and '*situated*' approaches to teaching" (Richards, 2008, p.167, italics added), video recorded lessons provided the necessary *visually* situated contexts. However, to mitigate the danger of video recorders having an intrusive effect on the context and observed events (Marshall & Rossman, 2006), I carried out a pilot prior to the main data collection using the recording equipment, in order to familiarise the teachers and students with it (see 4.5.2).

4.4.1.3 Audio recordings

The teachers were also audio-recorded throughout the observations, along with four student groups who were randomly chosen and participated on a voluntary basis. In order to minimise 'halo effects', as with the video recordings, a pilot was conducted (see 4.5.2).

4.4.1.4 Field notes and classroom materials

During the observations, I took unstructured field notes of what was going on in the classroom. This unstructured format allowed the phenomenon under investigation, that is, how the teachers used textbook tasks, to emerge without being constrained by pre-determined categories (Nunan, 1992; Nunan & Bailey, 2009). In addition to field notes, I also collected textbook unit documents and other classroom materials such as worksheets, pictures, powerpoint slides of the lessons, and so forth, to provide further contextual data. Since the study focussed on how teachers used textbook tasks in their classrooms and the rationales behind their task pedagogy, understanding of the process of what the teachers did with textbook tasks or of 'what was going on there?' allowed for a data-driven process in which the data guided the analysis (Mackey & Gass, 2005, p.179). In sum, the methodology allowed me to make sense of how teachers "theorize from their practice and practise what they theorize" (Kumaravadivelu, 2006a, p.173).

4.4.2 Stimulated recall and in-depth interviews with the teachers

According to Borg (2006), although observation is a useful technique for collecting data on teacher practices, it is by itself insufficient for investigating teachers' thoughts and beliefs. I therefore interviewed the teachers to derive the rationales for their practice of

textbook tasks. It is assumed that behaviours have their own purposes and express more submerged 'values and beliefs' (Marshall & Rossman, 2006). Interviews enable the researcher to obtain insights that cannot be inferred from observations (Mackey & Gass, 2005). They are appropriate to investigate "phenomena that are not directly observable" (Gass & Mackey, 2005, p.173), and "to make the invisible visible" (Kvale, 1996, p.53).

The interviews were carried out with each individual teacher in two ways: stimulated recall and in-depth interviews. The former were conducted as soon as practicable, sometimes immediately but not more than four days after the observed lessons and in separate sessions from the latter. The video recordings of the observed lessons were used as what Woods (1996) refers to as points of departure or contexts for the teachers to articulate the rationales for their task implementation (cf. Gass & Mackey, 2000). As Breen et al. (2001) comment, "we cannot deduce language pedagogies on the basis of teachers' accounts of how they work without reflecting with them upon *actual instances* of practice" (p.498, original emphasis). The procedures for conducting the stimulated recall interviews are described below.

- 1) I played selected parts of the videoed lesson(s) for the teachers to watch and asked them to recall what they did.
- 2) When the teachers said they were ready, I stopped the pause button and asked them to talk about the video clip.
- 3) I prompted them with further questions, especially questions pertaining to why certain choices were made.
- 4) I often drew to their attention links between their comments and textbook tasks, other classroom materials and the field notes. I did this to encourage the teachers to articulate and elaborate on their rationales.

The follow-up in-depth interviews further sought the teachers' elaboration and explanation on the issues identified in the earlier stimulated recall interviews. One concern raised by Borg (2006) is that teachers might articulate post-hoc rationales that might not be true of themselves. Addressing this concern, the teachers were asked to confirm and elaborate on what they said through "multiple-session format" interviews

(Dörnyei, 2007, p.135). In particular, each teacher took part in, on average, two stimulated recall interview sessions and one in-depth interview with me. Later sessions often went deeper into issues that had emerged in earlier ones and so consistent themes often appeared across lessons. This allowed me to cross-check information to confirm or disconfirm consistency. I also asked about issues beyond the observed instances of practice to further understand teachers' thoughts and beliefs (also see Andon & Eckerth, 2009). Overall, the multiple sessions assisted in checking what the teachers really meant and acted through "validation in situ" (Kvale & Brinkmann, 2009, p.249). The stimulated sessions lasted around one hour each and the in-depth sessions around 45 minutes.

The Vietnamese language was used in all these sessions to establish comfort and maximise understanding. A semi-structured format was used in which I worked loosely from "a written list of questions", and deviated or asked for more information when necessary (see Mackey & Gass, 2005, p.173). A semi-structured format was used for three reasons. First, it allows for prior preparation based on what has emerged from the observed lessons. Second, it allows for comparisons between the rationales of the individual teachers on their practices. Third, with a semi-open format, the interviewees are able to raise issues of concern or suggest new ideas (Weir & Roberts, 1994, p.146), further revealing the teachers' beliefs. Open questions were used in the interviews. I prepared the guiding questions based on what had been observed in the classroom, or in response to what I had found through the recorded lessons, field notes, and collected materials (see Appendix 4 for a sample of teacher interviews). All the interviews were audio recorded (see 4.5.2).

The purpose of the stimulated and in-depth interview sessions was for the teachers to recall, reflect and explain. As Cohen et al. (2007) point out, invalidity or bias can come from the interviewer seeking support for the concepts that he or she has perceived beforehand. To minimise this danger, I sought to respond to the interviewees' answers from a neutral stance "to provide an opportunity for reflection and further input" (Mackey & Gass, 2005, p.175). I did not make any comments that were coloured with evaluation or judgement of what the teachers said. In particular, I observed and practised neutral, supportive back-channelling using responses such as '*I see*', '*uh-huh*',

'ok', and so forth, to the teachers' answers. During the interviews I additionally encouraged the teachers to elaborate their answers by using '*Could you clarify that?*' '*Could you please talk more about this?*' '*So you mean?*' '*Why?*' '*Anything else?*'. Cohen et al. (2007) insist that interviewers must observe the principle of naturalness closely, that is, to elicit what is in the interviewers' mind without any attempts to intervene. I learnt from the piloting (see 4.5.2) that the teachers were put off by my taking notes while they were talking. For this reason, I audio recorded the interview (with the participants' permission) and took no notes. I concentrated on making face-to-face dialogue and after each interview, I wrote down reflections as soon as possible. As Fontana and Frey (2005) note, "each interview context is one of interaction and relation, and the result is as much a product of this social dynamic as it is the product of *accurate accounts and replies*" (p.699, italics added). The trust that I gained from the fact that the teachers were my peer colleagues (see 4.5.1), together with the use of Vietnamese L1 as the medium of the interviews, enabled me to establish an interview context where the interviewees were able to provide accounts that were true to themselves.

4.4.3 Stimulated recall focus group interviews with students

For student interviews, the focus group format (Bryman, 2008; Dörnyei, 2007) was used for several reasons. First, it made efficient use of limited time. Second, interviewees feel more comfortable and relaxed to state their viewpoints within a group (King & Horrocks, 2010). Furthermore, according to Marshall and Rossman (2006), when an interview involves more than one respondent, it affords opportunities for expansion and clarification and exploits "the collective experience of group brainstorming" (Dörnyei, 2007, p.144).

Fifty four students, six from each of the nine observed classes, voluntarily took part in focus group interview sessions after the classroom observations. Each session lasted around 1.5 hours. Like the teacher interviews, a semi-structured format was used. The guiding questions were pre-prepared, based on what was observed in the video recorded lessons, samples of which were used as stimuli. For example, I found that at the pre-task stage, some teachers provided a model of task performance before students rehearsed for their task performance. I therefore prepared the questions "How

do you like being provided some modelling like this (referring to the video)? And why?” Also noting that all the teachers followed a rehearsal-performance model, I prepared the questions “How do you like performing the task in front of the class?” and “Why?” Similarly, based on my initial analysis of the task rehearsals and performances that I collected, I found the students used Vietnamese L1 substantially. And so in the recall interviews, after students had viewed video clips of the rehearsal and performance stages, I raised this issue to elicit students’ viewpoints. In brief, based on what emerged in the data, I brought “increasing focus” to the questions of interest (Mackey & Gass, 2005, p.179) (see Appendix 5 for a sample of student interviews).

The student stimulated recall interviews followed the same technical procedures as the teacher stimulated recall interviews as outlined in 4.4.2. However, as the focus group interviews differed from the interviews with individual teachers, I further took the role of a moderator or facilitator (Bryman, 2008), who was there to invite and facilitate talk without interventions and tried to elicit ideas from all the students in the group. Importantly, I was not seeking agreement on viewpoints from all the students, but rather encouraged them to speak out their own individual ideas by inviting talk from students following up on what their friends had just said. I made it very clear at the beginning of the interview sessions that students did not need to agree with their friends, but were encouraged to give ideas that were true of themselves or of what they perceived.

4.4.4 Summary of data collection methods for Phase 1

A summary of the data collection methods used in the study is provided in Table 4.3.

TABLE 4.3: Summary of data collection methods

Methods	Participants	Time	Number of sessions
Classroom observations	9 teachers from Grade 10, 11, and 12	45 minutes/lesson	45 sessions, five per teacher
	Volunteer students being audio recorded for their task talk	four groups/lesson	180 audio recordings ⁸
Teacher stimulated recall and in-depth interviews	9 observed teachers	1 hour/stimulated session	18 sessions , two per teacher
		45 minutes/in-depth interview	9 sessions, one per teacher
Student focus group interviews	54 students, six from each of the nine observed classes	1.5 hours/session	9 sessions, one per class

4.5 Data collection procedures

4.5.1 Ethics

I obtained ethics approval from the Human Ethics Committee, Victoria University of Wellington before data collection began (see Appendix 1). Data were collected from 15 October-29 December, 2010. All the people involved (the Head of the school, the teachers, and students) were introduced to the research through verbal and written explanations and signed a consent form before the data collection began (see Appendix 2). I also sought the permission of the Director of the local Department of Training and Education before conducting the research at the high school.⁹ During the briefings with the Head of the school, the teachers and the students, I always emphasised that I was collecting data for the sole sake of research, and confidentiality and anonymity of the data were constantly guaranteed and they were entitled to withdraw from my research if they wanted. Furthermore, I encouraged the teachers and students to behave

⁸ However, only a small number of these were used due to technical problems (see 4.6.4.1).

⁹ In Vietnam, high schools are under general supervision and management of the local Department of Training and Education. I informed the Director of this Department of my research. She gave her consent and asked me to go and work directly with the school and the teachers and students.

normally as they do in their daily lessons. My position as a teacher and their colleague made it easy to observe without intrusion. I had never been in any position of power in relation with them; they treated me as a colleague friend, who was then doing research.

4.5.2 Piloting

Piloting started one week before the main data collection began. The purpose of the piloting was to familiarise the teachers and students with the equipment and habituate the teachers and students with my non-participant presence in the classroom to avoid the 'halo effects' (Mackey & Gass, 2005; McDonough & McDonough, 1997).

Furthermore, it also aimed to (1) test the feasibility of the recording equipment in capturing what was intended, (2) test the quality of the recordings, (3) supply data for sample-analysis, (4) try out interview procedures and (5) hone my data-collecting skills.

Two pilot observations were made to each of the nine classes. For each observation, I came to the class well in advance to set up the equipment before the bell rang. While the lesson was under way, I was sitting quietly in a corner at the back in the classroom behind a camera set up on a tripod, so that I was able to adjust the camera focus to the teacher moves and classroom events that occurred. Each classroom has free space on both sides, and the place where I sat with the camera was well at the back. I chose the place so that the camera was able to cover the largest number of students in the class. The teachers agreed to wear a small audio recorder on their shirt. I also put four digital audio recorders randomly among the (group of) students who were willing to be audio recorded. The teacher and students did not seem to pay attention to my presence in the classroom. In fact, during the first observation, at the beginning a few students occasionally played with the audio recorders put among themselves. However, on the subsequent trials, they did not seem to notice them. The video clips and my observations showed that the teachers and students did not seem to be affected by the equipment and that they were natural in their behaviour with much laughter and chatty responses. Furthermore, I also capitalised on breaks between classes to establish rapport by talking to students or involving them in helping me set up and collect the equipment. By the time the official observations began, the equipment and my presence

became part of the classroom life, the teachers and students appeared to behave normally.

During the piloting, in parts of the lessons that involved oral tasks, I learned that the teachers automatically asked the group (if they were called to perform the task in front of the class) to carry the recorder with them or students just did so without being requested. They were requested to do the same in the main data collection.

After the piloting with classroom observations, I also carried out brief interview trials with the teachers and the students with the aim of familiarizing them with the interviewing procedures. During the interviews (both stimulated recalls and follow-up interviews), I did not ask concrete questions that revealed the focus of the study. At first, I took notes during the interviews besides recording the conversations with a small audio recorder (with their permission). However, I found that taking notes impeded the conversation, and importantly the teachers and students seemed to feel that the conversation was getting more 'serious' and thus tended to speak more cautiously each time I noted down things they said. In the subsequent trials with other teachers and students, I did not take notes during the interviews, but audio-recorded them. The pilot interviews showed that the teachers and students acted naturally as if we were taking part in normal conversations or chatting about EFL teaching and learning in Vietnamese L1.

4.5.3 Main data collection

The main data collection began the week after the piloting. The same procedures as in the piloting were used. Due to the teachers' concurrent teaching time, in the first two weeks, I observed the Grade 10 and Grade 12 teachers, and in the following two weeks the Grade 11 teachers. Interview sessions were carried out as soon as practically possible after the observations (see 4.4.2).

4.5.4 Summary of data collection procedures

The procedures for data collection are summarised in Figure 4.1.

FIGURE 4.1: Summary of data collection procedures

Ethics approval from HEC, Victoria University of Wellington



Week 1	Recruiting teacher and student participants
Weeks 2-3	Piloting
Weeks 4-10	Main data collection

4.6 Data analysis

As the focus of the study was on oral tasks in the context of textbook unit teaching sequences, the analysis focussed on the lessons or parts of the lessons that involved oral tasks.

The data were inductively analysed through an iterative process of visiting and re-visiting the multiple data sources. Themes and categories were then derived from each teacher profile and all the teacher profiles. As the analysis of the qualitative data were theme-based, one associated criticism is that the reader is often left with the question of “the extent to which certain beliefs are held or a certain form of behaviour occurs” (Bryman, 2008, p.599). Bryman suggests one way to add greater accuracy to the findings is to quantify the data. This also finds support in other researchers who argue for the need to quantify qualitative data (Dörnyei, 2007; Ellis & Barkhuizen, 2005; Mackey & Gass, 2005; Nunan & Bailey, 2009; Silverman, 2010). According to Mackey and Gass (2005), quantification of qualitative data serves to (1) find patterns; (2) report data in a later stage; (3) explain why certain inferences are made; (4) provide a quick comparison to other findings in other settings. In this study, themes emerged and were informed by the data. They were then quantified to report the tendency or extent of ‘prevalence’ (Bryman, 2008).

I had the advantage of gaining accessibility and trust when collecting data at the high school where I had been teaching for 13 years. However, in my approach to observing, recording, and taking field notes, and later on, analysing the data, I constantly exercised ‘bracketing’ (Richards & Morse, 2007; Schostak, 2006). Specifically, I adopted “the role

of researcher as stranger, and thus see the familiar as strange ...” (Holliday, 2007, p.68). To elaborate, I focussed on recording, and describing what occurred and committed myself to ‘learning from’ and ‘making sense of’ the data (Richards, & Morse, 2007). The analysis concerning each of the research questions (set out at the beginning of the chapter) is specified below.

4.6.1 Analysing the teachers’ use of textbook tasks

4.6.1.1 How closely the teachers followed the textbooks

In order to answer this question, I first compiled a corpus of oral textbook tasks through 45 observed lessons that the teachers were supposed to use and the corresponding tasks that the teachers actually used in their classrooms. Examining and re-examining the data, I found three main ways in which the teachers used the textbook tasks: retained, adapted, and replaced, as defined in Table 4.4.

TABLE 4.4: Coding teachers’ use of oral tasks from the textbooks

Action	Definition
Retained	Teachers used exactly the same task as in the textbook
Adapted	Teacher used the textbook task, but modified it in various ways
Replaced	Teacher created own task, completely different from the textbook task

Upon reading and re-reading the various data sources, I summarised the teacher use of textbook tasks by devising a table which contained the oral textbook tasks in the taught unit in one column and the actual tasks or activities that the individual teachers of the same grade group used in one column. In this way I was able to see how the same textbook task was played out by different teachers. If the teacher task was exactly the same as the textbook task, I noted down ‘same’ and later coded as ‘retained’. If they were different, I wrote down the actual tasks in use and later analysed as ‘adapted’ or ‘replaced’ as defined above. Examples 1-3 show how I juxtaposed the textbook tasks and the actual tasks used by three different teachers.

Example 1:

Textbook task	Location	Teacher task	Teacher ID
Discussion based on the questions given: -Do you often go to a picnic? -What is the best time for a picnic? -Why do people go for picnics? (Unit 6: Excursion, English 10, p.67)	Pre-listening task	GAME: Teacher-led: Hidden picture: BOTANIC GARDEN (replaced)	110A
		GAME: Teacher-led: HANGMAN: Outdoor activities (replaced)	210B
		GAME: Teacher-led: Guessing picture-based activities (replaced)	310C

This analysis revealed that all the three Grade 10 teachers replaced the pre-listening textbook tasks with teacher-led games. I made similar summaries for the other two groups of teachers (Grade 11 and Grade 12) and a common pattern emerged.

For tasks in the other positions of the teaching sequence, I also devised similar tables. Example 2 shows that the three teachers replaced the textbook task in the post-reading stage with their own.

Example 2:

Textbook task	Location	Teacher task	Teacher ID
Completing a summary using the words given (Unit 6: Excursion, English 10, p.65)	Post-reading	Groupwork: Work in groups, discussing five essential things to bring for an overnight camping trip (replaced)	110A
		Groupwork: You have felt tired recently, and you want to go somewhere for relaxation (replaced)	210B
		Groupwork: Plan an excursion for some days off before Tet holidays (replaced)	310C

Example 3 shows how one textbook task in the speaking lesson was differently played out in the classroom by the three teachers.

Example 3:

Textbook tasks	Location	Teacher Tasks	Teacher ID
Groupwork: Work in groups. Read the seat plan and decide on the best seat for each participant using the participant information provided (Unit 6: Excursion, English 10, pp.66-67)	Speaking lesson	Groupwork: Work in groups of four, discussing what you are going to do in the next three days off (<i>replaced</i>)	110A
		<u>Same</u> (<i>retained</i>)	210B
		Groupwork: Plan your future excursion: where, when, what to do, how to get there and what to bring (<i>replaced</i>)	310C

By juxtaposing the teacher and textbook tasks, I was able to see how the textbook tasks were used by different teachers who taught the same lesson and detect patterns of use. The data were then tabulated into the frequency of retained, adapted and replaced tasks. From this stage of analysis, I found that all the nine teachers replaced the textbook tasks in the pre-reading and pre-listening lessons with teacher-led games, which made up one strand of divergence from the textbook. I therefore focussed on the actual tasks in use that the teachers adapted or replaced from the textbook tasks. I then looked for the task features where the textbook tasks and the teacher tasks differed. This is the focus of the next section.

4.6.1.2 Task design features

Table 4.5 summarises the features of task design that capture the changes the teachers made to the textbook tasks. Some of the themes were named using terms taken from the task-based literature.

TABLE 4.5: Features of task design modified by the teachers

Design feature		Definition
1. Input dependence	input-dependent	Task does/does not require students to depend on input provided (other than task instructions) to carry out the task.
	input-independent	
2. Goal orientation (Duff, 1986; Ellis, 2003)	convergent	Task does/does not require students to agree on a solution.
	divergent	
3. Solution type (Ellis, 2003)	closed	The task does/does not require a single, correct solution (or a limited range of correct solutions)
	open	
4. Personalisation	personalised	The task does/does not involve students talking about themselves, their lives, and their experiences.
	non-personalised	
5. Immediacy	(more) immediate	The task does/does not relate to the students' immediate world (immediate needs or interests).
	(more) remote	
6. Authenticity (Ellis, 2003)	(more) situational	The task does/does not involve a 'real-life situation'.
	(more) interactional	

These six task features represented an appropriate level of descriptive delicacy for capturing the nuances of teacher and textbook tasks. The first three features represent different layers of open-endedness that describe the nature of the changes made by the teachers in a way that a simple characterisation as open/closed tasks would not be able to do. For example, while closed tasks involve a single correct task solution, or a limited range of correct solutions (Long, 1985b), open tasks do not have any pre-determined correct solutions (Ellis, 2003). However, open tasks can differ in terms of the goal

orientation, whether it entails students to reach an agreement on the solution (convergent) or students can have their divergent viewpoints (divergent)(Duff, 1986). In fact, divergent and convergent tasks are sub-types of open tasks (Ellis, 2003, p.90). My data show that tasks can also differ in whether they involve students in using the input provided (other than task rubrics) in order to carry out the task or not (input-dependent and input-independent). For instance, a task that requires students to use the information provided about the participants who took part in a boat trip and the seat map to decide on the most suitable seat for each participant on the boat (see Example 6) is an input-dependent task because students rely on the input given to carry out the task. However, a task that requires students to discuss in groups what they are going to do in the next few days off is an input-independent task, as all the content comes from the students. Examples 4-5 show that an open task can be convergent or divergent, and similarly it can be input-dependent or input-independent, while a closed task is typically convergent and input-dependent (Example 6). This shows that all the three themes (input-independent/input-dependent, divergent/convergent, open/closed) are *necessary* to reveal the current data. And these three themes are straightforward and objective to code.

Example 4:

Work in groups, discussing five essential things to bring for an overnight camping trip (<i>Teacher speaking task, 110A</i>)	→	Open Convergent Input-independent
--	---	---

Example 5:

Work in pairs, discussing which piece of advice from the reading passage, do you think, the most useful (<i>Textbook post-reading task, Unit 6, English 12, p.65</i>)	→	Open Divergent Input-dependent
--	---	--

Example 6:

Work in groups. Read the seat plan and decide on the best seat for each participant using the participant information provided (Textbook speaking task 2, Unit 6, English 10, pp.66-67)	→	Closed
		Convergent
		Input-dependent

The remaining three design features from Table 4.5 denote different aspects of task authenticity in the current data that would not be captured solely by a label of situational or interactional authenticity defined by Ellis (2003). According to Ellis (2003), situational authenticity “concerns whether a task needs to correspond to some real world activity” (p.6) and it follows that not all tasks involve a real life situation. Tasks such as ‘describe and draw’ or ‘decide on who should stay in the plane’ would hardly ever occur in real life tasks in the target language, but “the kind of language behaviour they elicit corresponds to the kind of communicative behaviour that arises from performing real world tasks” (Ellis, 2003, p.6).

Ellis (2012, personal communication) posits that all tasks should aim to attain interactional authenticity, that is, to induce interactional processes similar to real life communication, and therefore tasks with situational authenticity are also tasks with interactional authenticity. Interactionally authentic tasks are not necessarily situationally authentic tasks. It seems to me, when I applied these concepts to my data, whether or not a task concerns situational authenticity is a matter of degree. That is, a task can be more situationally authentic than another because of its *likelihood* of occurring in real life in the target language. In light of the current data, I coded the teacher tasks and textbook tasks as more or less situationally authentic *in relation to each other*. However, in order to be consistent with the previous themes that involved binary variables such as closed/open, input-dependent/input-independent, and convergent/divergent, if a teacher or a textbook task is less situationally authentic than the other, they were coded as ‘more interactional’, even though ‘more interactional’ does not mean at all that the task involves students in processing language use of real life communication more deeply than the other. In two cases in the data (Examples 7 and 10) where both the

teacher task and the textbook task were equally situationally or interactionally authentic tasks, they were both coded as 'more situational' or 'more interactional'.

As the name indicates, Feature 4, personalisation, refers to whether students talked about themselves. Feature 5, immediacy¹⁰ concerns the context in which the task is situated, and whether it addresses students' immediate concerns or interests, as *explained by the teachers*. The following quote shows that immediacy arose from the teachers' words of reasoning, and it was needed to capture this theme in the current data.

*(1) Next week our students are going to have three days off because the provincial examination is going to be held at the school. So I **made use of the situation** and changed the textbook task to instead ask students to talk about what they are going to do during these days off. This is what students are **immediately interested in**, and they will engage more in the task. (110A)*

Although immediacy and personalisation (Feature 4) seem overlapping, they indicate different nuances that further distinguish the teacher tasks and the textbook tasks. Immediacy was needed to capture the difference between the teacher tasks and textbook tasks, especially in cases where both involved personalisation. For example, instead of using the textbook task that involved students talking about their future job (as to where, and who to work with, what salary to get, etc.), the teacher used a debate task where students have to defend their job choice. He said:

*(2) In real life, this issue (job choices) often causes tensions between parents and children of this age. I think my task would engage my students better, because this issue of debate reflects their **immediate concerns**. At this time point, Grade 12 students have to choose which university to make applications for and prepare for university entrance examinations, university of medicine, pharmacy, or pedagogy as these relate to their future jobs. They have to persuade their parents to let them choose their favourite university. Also students are still at school; they have never done any job yet; it is **too far***

¹⁰ In some studies (e.g., Foster & Skehan, 1996; Gilabert, 2007; Gilabert et al., 2009; Iwashita, MacNamara, & Elder, 2001; Robinson, 2001), immediacy refers to the contextual support of a task. For example, 'here and now' (with pictures) tasks are referred to as immediate and 'there and then' (without pictures) as remote (also see Ellis, 2003).

away to ask them to imagine the salary they will get, the working condition and so forth. (712G).

The phrases in bold reveal how the theme ‘immediacy’ came about in the data. In this case, although both the teacher task and textbook task involved students talking about themselves (their future jobs), the teacher task was coded as more immediate than the textbook task as reasoned by the teacher (see Example 7).

Example 7:

	TASK	Characteristics
Textbook	Work in groups, talking about your future job including -Where you will work -Who you will work with -The salary you may get -The working conditions <i>(Speaking task 3, unit 6, English 12, p.67)</i>	Personalised More remote More situational
Teacher	Debate: Work in groups, talking about your future job. You have to defend your job choice against your friends’ counter-arguments. <i>(712G)</i>	Personalised More immediate More situational

Again, for the sake of consistency, if a teacher task or a textbook task is less immediate than the other, it was coded as ‘more remote’. Of course in the case that the teacher task was personalised and the textbook task was not (or vice versa), the personalised task was synonymous with ‘more immediate’ and the non-personalised task was ‘more remote’ (see Example 8).

Example 8:

	TASK	Characteristics
Textbook	Work in groups. Read the seat plan and decide on the best seat for each participant using the participant information provided. (<i>Speaking task 2, Unit 6, English 10, pp.66-67</i>)	Non-personalised More remote More interactional
Teacher	Work in groups of four, discussing what you are going to do in the next 3 days off. (<i>110A</i>)	Personalised More immediate More situational

Some few examples (Example 9) where both teacher and textbook tasks were not personalised, but the teacher task involved simulations or role-plays and was coded as ‘more immediate’ because teachers perceived and presented them as such.

Example 9:

	TASK	Characteristics
Textbook	Groupwork: Discuss which piece of advice given in the passage you find most useful and least useful. Why? (<i>Post-reading task, unit 6, English 12, p.65</i>)	Non-personalised More remote More interactional
Teacher	Role-play: Student A plays a job candidate and Student B plays a teacher Job candidate: seeks the teacher’s advice on how to prepare for a job interview (<i>912I</i>)	Non-personalised More immediate More situational

The teacher reasoned that the textbook task was more ‘far away’ than her task because in her task, students had a chance to practice their future situation.

(3) Here students can give their own advice, so it was **closer to them** and the situation is likely that in the future when they will be seeking advice to attend a job interview. (912I)

As in the case of Feature 6, there were three cases where both the textbook task and the teacher task were coded as ‘more remote’ because neither of the tasks was ‘more immediate’ than the other (Example 10).

Example 10:

	TASK	Characteristics
Textbook	Pairwork: Rank the following causes of overpopulation in order of importance and explain why <i>(Speaking task 1, unit 7, English 11, p.83)</i>	Non-personalised More remote More interactional
Teacher	Groupwork: Discuss causes of overpopulation <i>(611F)</i>	Non-personalised More remote More interactional

I coded the textbook tasks and the teacher tasks for the first three features (Table 4.5) first, and then coded them for the last three features. Then I summarised coding tables of each pair of textbook task and teacher task as in Example 11. I subsequently calculated the frequency of each of the six themes for the textbook tasks and the teacher tasks to detect the patterns. Layered and focussed coding enhanced accuracy (Révész, 2012).

Example 11:

	TASK	Characteristics
Textbook	Work in groups. Read the seat plan and decide on the best seat for each participant using the participant information provided. (<i>Speaking task 2, Unit 6, English 10, pp.66-67</i>)	Input-dependent Convergent Closed Non-personalised More remote More interactional
Teacher	Work in groups of four, discussing what you are going to do in the next 3 days off. (<i>110A</i>)	Input-independent Divergent Open Personalised More immediate More situational

A Vietnamese EFL teacher was trained to independently code the textbook tasks and the teacher tasks using the six design features in Table 4.5 above. She first coded these tasks for the first three features, and then coded them for the last three features. The Cohen Kappa (κ) was .96 for teacher tasks and .95 for textbook tasks, indicating high inter-rater reliability. Teachers' task rationales were also provided to the second coder for coding Feature 5.

4.6.2 Analysing the teachers' procedures of task implementation

The observational data showed that all the teachers organised their task-based lesson into four distinct phases: pre-task, rehearsal, performance and post-task. Analysis involved describing exactly what each individual teacher did at each phase of their task implementation by examination and re-examination of the multiple data sources and noted frequency of certain practices to find patterns for each teacher and all the teachers. Observed practices were later on reflected on and explained by the teachers in the stimulated recalls and interviews.

4.6.3 Analysing the interview data

I transcribed all the interviews (both teacher and student) in their entirety and then translated the material into English. Care was taken to maintain the “accuracy and subtlety in translation” (Marshall & Rossman, 2006, p.112). I tried to retain the closest meaning to the original when translating. When I was not sure about English equivalents for certain words that the teachers used, I kept the L1 words that the teachers used in brackets for later cross-checking and refinement. The data were then double-checked for the accuracy of transcriptions and translations by another Vietnamese EFL teacher before analysis.

Analysis of the interview data was also an iterative process. I inductively looked for the ‘sounds’ or ‘flavour’ in the data in order to ‘learn from’ and ‘make sense of’ them (Richards, 2009; Richards, & Morse, 2007). It was a process of “*sifting, or combing or searching*” (Nunan & Bailey, 2009, p.416), for which I followed the five steps proposed by Lieblich, Tuval-Mashiach and Zilber (1998) and recommended by Nunan and Bailey (2009, pp.423-424):

- 1) I read the interview data multiple times to find a pattern. Finding a pattern involves searching for i) “repeated themes or key words”; ii) “parallel or connected comments” or iii) “metaphoric use of language” that carries a salient message (Nunan & Bailey, 2009, pp.416-417).
- 2) I wrote down my overall initial feelings.
- 3) I decided on the ‘content’ or themes that emerged as the data unfolded “from beginning to end.”
- 4) I highlighted the themes that emerged and read them “separately and repeatedly for each one”.
- 5) I followed each of the themes throughout the data and made some ‘conclusions’. Here I was also alert to the appearance of each theme as to the setting where it occurred and how ‘salient’ it was in the data.

When I was searching for themes, I treated each theme instance that transpired as ‘provisional knowledge’ that needs further examination for confirmation (Silverman,

2010, p.279). In other words, I tested and re-tested themes that surfaced through comparison and scrutinisation of the data for each teacher and all the teachers. Furthermore, following Richards (2009), I established mutual hyperlinks in the data for each teacher and to other teachers, which allowed me to readily navigate and check and double-check any claim before it was made. Furthermore, the Vietnamese and the English versions were always available each time I did the analysis, facilitating clarification if needed. Thus researcher subjectivity and bias was minimised. I also used the teachers' own words as one of the sources of validation. As described earlier, I used quantification to detect and report patterns. The semi-structured interview format (see 4.4.2) enabled me to compare what one teacher/student said about certain practices with another teacher's/another student's reported perceptions. In this way, I was also able to find the frequency for certain beliefs held behind certain practices.

For example, in order to answer why they diverged from the textbook tasks, the interview data were searched for statements of rationale. This was first done for each individual case teacher before all the cases were brought together to gain an overall picture of tendencies in the teachers' personal theories of designing and managing oral tasks. Certain repeated words in the data revealed a socio-affective dimension to the teachers' descriptions such as "dry", "whole-heartedly", "boring", "hot", "worthwhile", and "engagingly". Socio-affective engagement thus emerged as a central theme that guided the teachers' task pedagogy. Another example involved the teacher rationales for public performance. When asked why they used public performance that followed rehearsal, although the teachers used different words/expressions such as performance was "*the climax of the task*" (210B), "*the happy ending*" (712G), or "*(without performance), I feel like there is something lacking, like the lesson is incomplete*" (310C), these expressions denoted one common theme, the 'happy ending' of the lesson. "Happy ending" was thus one theme that emerged and accounted for the teachers' use of public performance.

Analysing the student interview data followed the same approach for analysing the teacher interview data, as described above.

4.6.4 Analysing the learners' task-based interaction data

4.6.4.1 The data set

Recall that I put 4 audio recorders among volunteer students to capture their talk throughout the 45 observations, although my focus was only on oral tasks. All together I collected 180 transcripts of learner-learner interaction. Due to background noise and other extraneous factors, some data were discarded, and the final data set for analysis included forty-eight (48) transcripts of student task rehearsals in pairs/groups collected during the 45 observed lessons, across the 9 classes. It is of note here that the data were gathered in different positions of the lessons where the teachers used oral tasks. The transcribed rehearsals totalled 247.97 minutes, ranging from 4.08 to 9.07 minutes ($M=5.16$; $SD=1.09$). However, the number of performances by the same group whose rehearsals had been recorded was very limited (15), since the teachers randomly called on students to perform the task in front of the class.

The data involved students doing mostly open-ended tasks that the teachers created themselves or adapted from the textbook as will be shown in Chapter 5. Some examples of these tasks are:

- Work in pairs, discussing a plan for a picnic for the next three days off.
- Work in pairs, discussing your future jobs, why you like that job, and why.
- Work in groups, discussing five essential things to bring on an overnight camping trip.
- Work in groups, discussing causes of overpopulation.
- Work in pairs, doing a role-play between a person who seeks advice to attend a job interview and an advisor who gives advice.

4.6.4.2 Coding LREs in rehearsal and uptake in performance

All the transcribed rehearsals were coded for LREs, which, according to Swain (1998), are “any part of a dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct” (p.70) (see details in Chapter 7). As the focus was on learner-initiated LREs, pieces of the data that contained episodes

where the teacher responded to students' errors, or initiated focus on form were not included for analysis. However, the LREs were extended to incorporate episodes where students occasionally requested assistance from the teacher (also see Williams, 1999).

The 15 performances were coded for whether the language points that had been attended to in LREs in the equivalent task rehearsals were taken up correctly (see successful uptake in Chapter 7). These preliminary findings provided the underpinnings for the Phase 2 study of the thesis.

Another Vietnamese teacher of English was trained to code randomly 25% of the rehearsal data (12 task rehearsals) for LREs, and the agreement percentage was 89%. This same teacher also coded the 15 task performance for uptake, and the simple percentage agreement was 94% (see Chapter 7 for how simple percentage agreement was calculated). All the disagreements were resolved through a follow-up discussion. Those items where both coders could not reach an agreement on were excluded from analysis. The final agreement rates were 92% for LREs and 97% for uptake.

4.6.4.3 Coding amounts of L1 and L2 use

Because a turn might contain a long sentence or just one single word, both turns and words were used to analyse the amount of language use in Vietnamese L1 and English L2 in task rehearsal (also see Storch & Aldosari, 2010).

Coding language turns was a straightforward and objective process. Turns and words "are easy to identify, and this means high coding reliability" (Storch & Aldosari, 2010, p.361). In this study, turns in learner oral production were bounded by what each interlocutor spoke at one time. Each transcribed rehearsal with student speakers identified (S1, S2, S3, etc.) was then looked for L1 turns, L2 turns and mixed language turns. An L1 turn was the turn where Vietnamese L1 was entirely used. Similarly, L2 (English) turns were turns where English was completely used. Mixed language turns were turns where both languages were used, even if one was used more than the other. After separating turns into these three types, turns of each type were copied and pasted into Microsoft Excel sheets to calculate the frequency (by counting the rows).

To measure the amount of L1 words, L1 words were separated from the total words (L1+L2) in each task rehearsal, and counted for frequency using the word count function in the Microsoft Word (also see Storch & Aldosari, 2010). The frequency data were then tabulated and compared using statistical tests as appropriate.

The functions of L1 were exemplified through transcripts and students' L1 interviews data.

4.7 Validity and reliability

As already stated, I preferred to call this Phase 1 study a 'descriptive' study (Ellis, 2012). I analysed data inductively or qualitatively to look for themes that emerged, and then I quantified the data to find patterns or tendencies for certain practices and beliefs, allowing for inter-coder reliability. Some researchers use different labels such as 'credibility' and 'dependability' to alternatively refer to validity and reliability in research that involves qualitative data, for example from classroom observations and interviews (Bryman, 2008; Mackey & Gass, 2005). With both qualitative and quantitative approaches to data analysis, I have used the terms 'validity' and 'reliability' to discuss the 'trustworthiness' of the study. This section provides a stocktaking of my endeavours to enhance the validity and reliability of the study.

4.7.1 Validity

The validity of a study refers to how accurate the results are in achieving what it sets out to investigate (Creswell, 2009; Mackey & Gass, 2005). The current study involved 'qualitative' data such as classroom observation and interview data. Thus it might be criticised for *subjectivity*, or *lack of transparency* in analysis (Bryman, 2008). Subjectivity is an issue because the researcher is the 'instrument' (Kvale & Brinkmann, 2009; Richards & Morse, 2007) or the researcher "must make sense of the data" (Holliday, 2007) and find patterns because "themes and patterns do not simply jump out at the researcher" (Ellis & Barkhuizen, 2005, p.525). However, validity of this kind of 'qualitative research', according to Kvale and Brinkmann (2009), pertains to "validity as quality of the craftsmanship" (p.248). This quality of craftsmanship was demonstrated through the measures the study took below:

- (1) A research design that took into account a fit between research questions and data sources.
- (2) Purposive sampling of the teachers and classes (i.e., teachers and students who used task-based textbooks).
- (3) Multiple case studies (nine teachers doing teaching sequences; three from each grade levels, teaching the same textbook units), thus enhancing the consistency and accuracy of the findings.
- (4) Piloting to familiarise the participants with the recording equipment, the procedures, and the presence of the researcher in the classroom to prevent 'halo effects'.
- (5) The researcher's 'bracketing' in data collection and analysis to minimise subjectivity.
- (6) A repeated data-driven process of analysis.
- (7) A semi-structured interview format that allowed for cross-checking the data among the teachers/students and finding patterns.
- (8) 'Multiple-session' interviews with the teachers to confirm consistency and accuracy in what the teachers said to prevent 'post-hoc' rationales.
- (9) Quantification of the data, which did justice to the validity of the patterns found, thus precluding criticisms of 'anecdotalism', bias caused by only a few demonstrations of the phenomenon that the researcher picks from the data (Silverman, 2010).
- (10) Multiple data sources such as triangulation to examine i) how the teachers used and implemented textbook tasks (classroom observations: audio/video recordings, field notes, textbook and classroom materials); ii) the teachers' rationales (teacher stimulated recall and in-depth interviews); and iii) how learners engaged in classroom tasks (video and audio recordings of student task-based interaction) and students' perceptions (student focus group interviews).
- (11) A holistic perspective on tasks in action which tracked tasks from textbook tasks to teacher tasks, and from teacher tasks to learner tasks, thus enhancing the ecological validity of the research.

4.7.2 Reliability

The reliability of a study addresses how replicable the results are, that is, whether the same results will be obtained by another researcher using the same methods (Creswell, 2009). The current study achieved reliability standards through:

- (1) Explicit description of the research context, the research participants and data collection methods.
- (2) Explicit and transparent analysis of the data: step by step coding of the teachers' use of textbooks, juxtaposition of textbook and teacher tasks, and finding patterns of task design features and task implementation procedures.
- (3) Focussed coding of the learners' interaction data and coding protocols: LREs and uptake, amounts of L1 and L2 use with explicit coding protocols.
- (4) Quantification of the data not only made data analysis transparent. It would also enable quick comparisons to be made of findings from other studies in different contexts (Mackey & Gass, 2005).
- (5) Formal inter-reliability tests that yielded satisfactory standards.

4.8 Summary

The aim of Phase 1 was to describe how the teachers actually used and implemented the oral textbook tasks in their classrooms, and how the students were engaged in the tasks that the teachers provided. The study adopted a descriptive research paradigm that aligned with its purposes. The data sources included classroom observations with field notes, video and audio recordings and transcripts of teacher discourse, textbook and classroom materials, teacher and student interviews, and transcripts of student task-based interaction. The data were analysed both qualitatively and quantitatively. The study took into consideration the issues of validity and reliability and took careful steps to enhance them.

The next chapter (Chapter 5) will present and discuss the findings on how the teachers used the textbook tasks and their ways of implementing tasks in the classroom. The

findings on how the learners engaged in the classroom tasks and their perceptions of learning opportunities through tasks will be covered in Chapter 6.

Chapter 5 THE PHASE 1 STUDY: HOW THE TEACHERS IMPLEMENTED TEXTBOOK TASKS

5.1 Introduction

The Phase1 study investigated the teachers' use of the oral tasks in the prescribed textbooks, including task design features and implementation procedures. In this chapter¹¹, I will present and discuss the findings related to each of these research questions (RQs) that the study aimed to answer:

- RQ1. How closely did the teachers follow the textbook tasks?
- RQ2. In what ways did the design features of the teachers' tasks diverge from the textbook tasks, and why did they diverge from them?
- RQ3. What task implementation procedures did the teachers use, and why did they use them?

5.2 How closely did the teachers follow the textbook tasks?

All the observations were of a sequence of five lessons of a textbook unit for each of the nine teachers. Each unit contained five lessons: Reading, Speaking, Listening, Writing and Language Focus. In addition to the speaking lessons, oral tasks are present in the pre- and post- reading and pre- and post-listening sessions for the reading and listening lessons. No oral tasks are prescribed in the writing and language focus lessons of the unit.

The data (Table 5.1) show that there was a strong tendency for the teachers to create their own tasks or adapt the textbook tasks. Out of the 64 oral tasks, 43 replaced the original textbook tasks, 12 were adapted from the textbook tasks and only nine were tasks as they were presented in the textbook. The tasks that were adapted or created by the teachers thus account for 86% (55/64) of all tasks. This tendency to diverge from the textbook was consistent for the teachers at all three grade levels.

¹¹ This part has been presented at the 4th Biennial International Task-Based Language Teaching Conference at the University of Auckland, 18-20 November, 2011, and has been submitted for publication (Publisher TBA).

TABLE 5.1: Teacher use of (oral) textbook tasks

Textbook task use	Grade 10 (n=3)	Grade 11 (n=3)	Grade 12 (n = 3)	Total (n = 9)
Retained	0	5	4	9
Adapted	2	5	5	12
Replaced	19	9	15	43
<i>Total</i>	<i>21</i>	<i>19</i>	<i>24</i>	<i>64</i>

Note. n= number of teachers

5.3 In what ways did the design features of the teachers' tasks diverge from the textbook tasks, and why did they diverge from them?

5.3.1 In what ways did the design features of the teachers' tasks diverge from the textbook tasks?

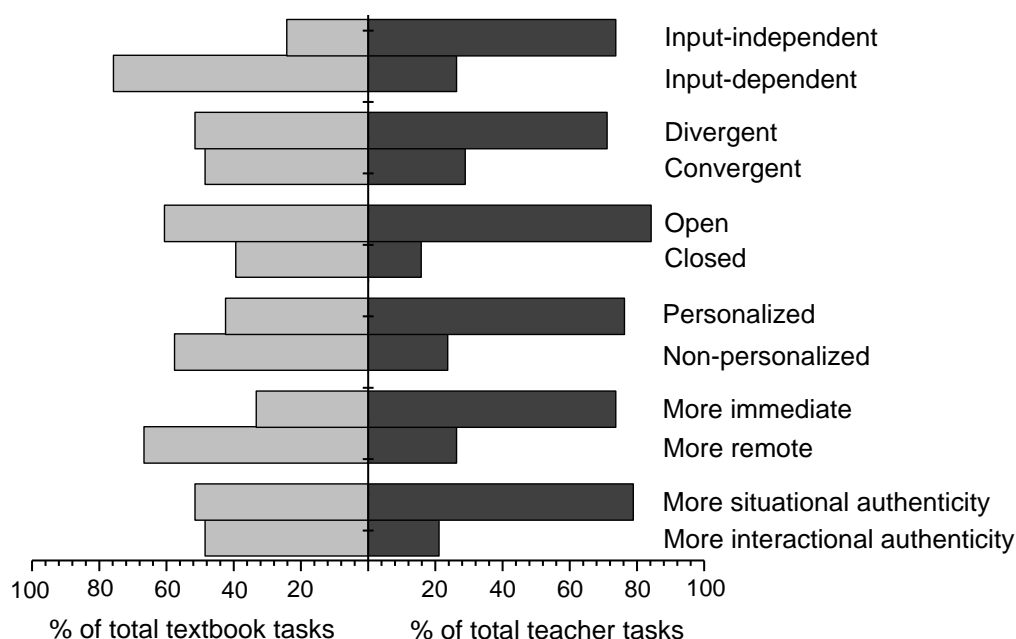
For the pre-reading and pre-listening oral tasks, the nine teachers replaced 18 textbook tasks, nine in the pre-reading and nine in pre-listening sessions. In all but one they were replaced with games followed by brief teacher-led question-answer sessions. Only one Grade 11 teacher used her own oral task instead of the oral pre-listening textbook task. This constituted a strand of divergence in its own right. The teachers stated the underlying rationale behind this practice was the 'fit for purpose' principle that the teachers adopted. They reasoned that these teacher-led activities were effective in leading into the listening/reading passage, and they fitted the purpose of the reading and listening lessons. In this section, the focus will be therefore on the task design features of the remaining 38 of the 55 oral communicative tasks that the teachers either adapted or created themselves, particularly how they differed from the textbook tasks. The task design features of the textbook tasks and the teachers' tasks are presented in Table 5.2 (also see Table 4.5).

TABLE 5.2: Design features of textbook and teacher tasks

	Feature	Textbook task		Teacher task		χ^2	<i>p</i> (* <i>p</i> < .05)
		n	%	n	%		
1	Input-independent	8	24.2	28	73.7	17.273	.000*
	Input-dependent	25	75.8	10	26.3		
2	Divergent	17	51.5	27	71.1	2.861	.091
	Convergent	16	48.5	11	28.9		
3	Open	20	60.6	32	84.2	5.021	.025*
	Closed	13	39.4	6	15.8		
4	Personalised	14	42.4	29	76.3	8.494	.004*
	Non-personalised	19	57.6	9	23.7		
5	More immediate	11	33.3	28	73.7	11.616	.001*
	More remote	22	66.7	10	26.3		
6	More situational (authenticity)	16	48.5	30	78.9	7.184	.007*
	More interactional (authenticity)	17	51.5	8	21.1		

Table 5.2 shows that for each feature the tasks designed by the teachers diverged from those prescribed in the textbook. For example, in the case of input-dependence, eight of the textbook tasks were input-independent and 25 input-dependent whereas the teacher-designed tasks were the mirror image of this with 28 input-independent tasks and 10 input-dependent tasks. Chi-square analysis shows significant differences ($p < .05$) between teacher and textbook tasks for a majority of the task features, except Feature 2. Despite the lack of significant difference, a closer look at the teacher tasks for Feature 2 reveals that divergent tasks (27) are more than double the number of convergent tasks (11). Figure 5.1 provides a more visual representation of the differences between teacher and textbook tasks.

FIGURE 5.1: Design features of textbook and teacher tasks



As shown in Figure 5.1, for most of the six task features, the teacher tasks showed a stronger tendency to be of one kind rather than another (e.g., input-independent rather than input-dependent, divergent rather than convergent, open rather than closed), while the textbook tasks were less predictably of either kind. The teachers showed a strong preference not only for tasks with more situational authenticity in general, but also for tasks that were personalised and concerned with the students' immediate world. Interestingly, although each of the three teachers who taught the same textbook units at each grade level used different tasks, they shared a 'collective practice' (Breen et al., 2001) in the way in which they diverged from the textbook task design. It is reasonable to assume that underlying this collective practice is decision-making that has something in common across all the teachers. The next section explores this decision-making.

5.3.2 Why did the teachers diverge from the textbook tasks?

The next question, then, is *why* the teachers diverged from oral textbook tasks in the ways they did. What emerged strongly from the teachers' rationales in talking about choices of task was the socio-affective dimension. This can be illustrated with a selection

of the words that were used by the teachers in talking about their preference for open-ended tasks (open, divergent, and input-independent) and tasks with situational authenticity, particularly tasks that were personalised to address students' immediacy as explained in Chapter 4. Table 5.3 shows the socio-affective flavour of the nine teachers' stimulated recalls and in-depth interviews.

TABLE 5.3: The socio-affective flavour in the nine teachers' talking about tasks

Teacher ID	The socio-affective dimension in teachers' talking about tasks
110A	Fun, interesting, convenient, relaxing, flexible, eager, boring, graceless, enthusiastic, meaningless, enjoyable, freer, worthwhile, motivating, close to student life, 'hot', instant, immediate, sad, creative, understanding, real
210B	Dry, poor, heavy, monotonous, relaxing, enjoyable, sad, practical, too specific, free, new, refreshing, passive, active, creative, cooperative, relevant, (un)willing, fun; student expectations, interests, likes and dislikes
310C	Confusing, dry, complex, tiring, too heavy, boring, suitable, relevant, engagingly, wholeheartedly; student likes, dislikes, student psychology, emotional feelings, interests, age, classroom atmosphere, student interests, student life
411D	Happy, enjoyable, enthusiastic, exciting, motivated; more robust, powerful, gorgeous, appealing, merrier, graceless, non-sense, worthwhile, more responsible, passive, lazy, sad, productive, creative, disappointing, freely; constraining, dry, boring, engaged; classroom atmosphere, student interests
511E	More suitable, easier, exciting atmosphere, time-consuming, more interesting; enthusiasm, interest, more reticence, lacking something, established routine, happy ending, classroom atmosphere; student interests, student life
611F	Collaborative, enjoyable, interesting, dynamic, necessary, safe, comfortable, distant, close; learner levels, learner expectations

712G	Boring, interesting, lively, logical, relevant, worthwhile, actively, not passively, close, far away (to student life), happier, more delighted, light, enjoyable, natural, stressed, heavy, fun; repetition and boredom, classroom atmosphere; immediate concerns; student interests, age, student life, personal development
812H	Real world, realistic, free, welcome, encouraging, fast, suitable, effective, more interesting, relevant, close to student life, relevant to exams, fun, warm-up, ice-breaking, already available, new, change, free, confident, motivating, enjoyable, enthusiastic; 'far away'; classroom atmosphere, students' boredom, shyness
912I	Motivating, interesting, enthusiastic, cooperative, real life, engaging, artificial; learner motivation and interests, expectations, student personality

The actual words the teachers used in Table 5.3, though de-contextualised, represent a strong socio-affective dimension in the teachers' comments. The following sections provide further detail of this dimension in the teachers' practical theory of oral tasks in relation to their classroom practice, focussing on open-ended tasks (open, divergent, input-independent) and authentic tasks (situational authenticity, personalised, and immediate).

5.3.2.1 Open-ended tasks

In talking about the underlying grounds for choosing open-ended tasks, the teachers articulated definite beliefs about their expectations of how students would respond to the task. For instance:

- (1) *This [textbook] task is constraining, dry and boring, as it requires students to rank these causes of overpopulation in order of importance. Students can think of causes of overpopulation themselves, they can think of many more ideas than those provided. And why to rank? There is nothing much in there. If I ask them to do this, they will do it unwillingly, not much language use; students' use of L1 is unavoidable, so I must design tasks that engage them, so that they are motivated to talk. Students like tasks that allow them to think for themselves, as they are very creative, you know. If they can speak freely, they will be more engaged and speak more with interesting ideas. (411D)*

- (2) *This textbook task? Boring and graceless! I used this task before, students were not interested, not eager, not enthusiastic; the atmosphere was sad; students only stood up and talked about seats for certain people on the boat, not much language produced; students used Vietnamese to complete the task as quickly as possible. Students don't like it; they like something more enjoyable, freer. If given a chance to talk about what they are going to do in the next three days off, students will search for more words, they will talk more, talk better and it will be more enjoyable. Here discussing the seats on a boat, students will soon find the solutions; the language output will be limited, only A should sit here, B there, simple language, boring!(110A)*

The primary teacher concern in these and other extracts is that they were aware of the socio-affective impact when tasks do not engage the students. Such words as 'constraining', 'dry', 'motivated', 'boring', 'graceless', 'eager', 'enthusiastic', 'sad', and 'enjoyable', convey the dominance of the socio-affective dimension in the teachers' thinking about oral tasks. The actual engagement, according to the teachers, is dependent on the degree to which tasks allow students to enjoy the task work. This need for maintaining creative engagement is the key reason given for the choice of open-ended tasks.

Underlying the need for engagement is the requirement of the teachers that the learners extend themselves in their use of English. The teachers were concerned that goal-convergent tasks tended to induce students to use simple language and L1 to jump to the outcome of the task at the expense of English use. Research shows that in EFL contexts, students resort to L1 use when doing pair/group work (Algería de la Colina & García Mayo, 2009; Guk & Kellogg, 2007; Storch & Aldorsari, 2010) and the teachers in the current study were well aware of students' reliance on the mother tongue (as also documented in Bock, 2000; Carless, 2004, 2008; Pham, 2007; Lee, 2005). The *quantity* of English used by learners figured highly in the teachers' beliefs as to what was desirable. Maximizing output was explicitly made reference to, and reiterated by all the nine teachers, as the ultimate outcomes of oral tasks, as illustrated in the following comments:

- (3) *Whatever you do, the ultimate aim is language output, as long as students produce a lot of output in English, the task is a success. So tasks should engage students; students have to like the task first in order for them to produce language output. Open discussions or conversations work, as students like them more than constraining tasks. (812H)*
- (4) *What I am most concerned about is how to choose a task that induces as much talk from students as possible. Open tasks like free discussions or role-plays that engage students both socially and emotionally will motivate them to talk. (611F)*

Thus, in summary, the teachers held a strong view that students need to be first engaged by the task socio-affectively in order to produce as much output as possible. In this regard, they expressed belief in the positive value of open-ended tasks. This belief resonates with earlier criticisms of task-based learning concerning the quantity and quality of output (Seedhouse, 1999). In particular, Seedhouse criticised closed tasks for generating simplified or ‘minimalised’ and ‘indexicalised’ language. In line with this view, the teachers, drawing on the pragmatic understanding of students’ likes and dislikes and their own EFL situation, sought open-ended tasks as they saw the potential of this task type in engaging students and generating more language output. This also aligned with Ellis’s (2003) recommendation that greater varied language use can result if open divergent tasks are used in combination with pre-task planning and teacher monitoring. In this regard, the teachers used open-ended tasks, but in a rehearsal-performance condition where in the former the learners were given some time to prepare for the task to go public (performance) (see 5.3). Further implied in earlier comments (Comment 2) was a link between open-ended tasks and the English output in the public task performance. As teacher 110A commented, “the atmosphere was sad; students only *stood up* and talked about seats for certain people on the boat, not much language produced Only A should sit here, B there, simple language, boring!” In this respect, the teachers’ preference for open tasks seemed to be linked to the space these tasks had for public performance.

The value that these teachers put on open-ended tasks is in conflict with the general claim that open tasks supply fewer opportunities for learners to develop their interlanguage than closed convergent tasks (Long, 1983, 1996; Pica, 1994; Pica et al.,

1993; Pica, Lincoln-Porter, Paninos, & Linnell, 1996)(see Ellis, 2003; Mackey, 2012 for recent reviews). Interestingly, Nakahama et al. (2001) found that the open task, that is, the ‘conversational activity’ in their study, had greater potential for rich multi-layered interaction, although it generated fewer instances of negotiation of meaning. Clearly, while the potential of open-ended tasks needs further investigation, it might well be time to look beyond instances of negotiation of meaning as indices of the values of tasks, an argument proposed by Foster and Ohta (2005) and Nakahama et al. (2001). In addition, while a specific non-linguistic outcome is an essential part of task definitions (e.g., Bygate et al., 2001; Ellis, 2003, 2009a; Skehan, 1998; Van den Branden, 2006a), the teachers emphasised the process of doing the task (cf. Andon & Eckerth, 2009; Prahbu, 1987), the socio-affective engagement it generates and the need to generate opportunities for language output in L2. These views are not in opposition to a definition of tasks as having a non-linguistic outcome because this feature is intended to foster language use, but it does suggest that teacher decision-making that is more focussed on how well language use is fostered in a specific context is a valuable object of enquiry. It would be useful to remind ourselves that a task serves “the overall aim of promoting language learning, through *process or product or both*” (Samuda & Bygate, 2008, p.69, italics added). It is therefore worth obtaining the data on learners’ performance in open as opposed to closed tasks or divergent as opposed to convergent tasks to see whether the teachers’ thinking and their practice is reflected in the learners’ task work. This was the reason for the inclusion of these two latter tasks in the Phase 2 study of the thesis.

In accord with these teachers’ practice and belief in the potential of open-ended tasks, the fact that they preferred tasks with situational authenticity, and in particular, tasks that are ‘real’ to students (personalised, with greater immediacy), appeared to fit in quite naturally with the socio-affective engagement line of argument. I now turn to these aspects of task authenticity to further understand the teachers’ practical theory.

5.3.2.2 Authentic tasks

The real world application of oral tasks (situational authenticity) was an important criterion in teachers’ rationales. Oral tasks, according to them, should allow for application in the real world when students leave school. Seven out of the nine teachers

made explicit reference to this aspect of oral tasks. The following comments are representative:

- (5) *Later in life students should be able to speak, to communicate in English in their future jobs or future studies. The real world application can motivate students to work. (511E)*
- (6) *I often adapt textbook tasks or use my own tasks. I tend not to use tasks that are not practical, not 'real world'. What I pay particular attention to is the real world application of the task into students' future life. For example, in this lesson, instead of using the textbook task, which involves students working out a summary of the listening passage in pairs, I asked students to do role-playing, that is, one is a job candidate, and the other an employer. Students will be able to experience what a job interview is like. I want students to prepare this in advance because later in their life, they will have to apply for a job any way, for example in a foreign company, this experience is preparing them for future use of English. Practical tasks engage students more than 'far away' tasks. (812H)*

Teacher perceptions of the need for a practical or real world feature of oral tasks make sense in an EFL teaching environment where students might not see the relevance of communicative tasks (e.g., Bock, 2000; McDonough, 2004; Pham, 2007), unless they are obviously related to their future jobs or future studies or pursuits (McDonough & Chaikitmongkol, 2007). Unlike the four teachers in Andon and Eckerth's (2009) ESL study, who used tasks with both situational and interactional authenticity, the teachers in the current study showed a strong preference for tasks with situational authenticity (rather than interactional authenticity).

In the search for engagement, the teachers also emphasised relevance through personalizing the tasks, connecting them to students' current lives. The nine case teachers referred frequently to this aspect in the interviews:

- (7) *Tasks, importantly, should be real to students; students are not going to engage themselves in any task you ask them to do, but they are in tasks that are relevant to them.... In real life, this issue (job choices) often causes tensions between*

parents and children of this age. I think my own task would engage my students better, because this issue of debate reflects their immediate concerns. (712G)

- (8) *This task does not make good sense to students. Although students have a chance to talk about themselves, this second task, deciding on the best seats for themselves on a given boat plan is not worth discussing. First, students rarely go by boat. Second, our students usually don't care about where to sit, except maybe some seasick students, but students of this kind, not many, may be one or two, so it's not real to ask students to do this task. Because if you ask them to, they still do it, but not engagingly or wholeheartedly, I assume. (310C)*
- (9) *Confronted with each textbook topic and associated tasks, I always consider whether I can make it better for my students or not and how. Personalizing the tasks, bringing it closer to student life is what I often do. How students do a given task all depends on whether it is relevant or interesting to them or not. (210B)*

These comments reflect the view (Samuda & Bygate, 2008) that tasks need to address “personal relevance and learners’ own experience as catalysts for learning” (p.221) (also see Willis & Willis, 2007). The teachers’ choice of tasks emphasises learner motivation and engagement, which is central to Dogme principles in ELT teaching (Meddings & Thornbury, 2003; Thornbury, 2012). Interestingly, while Long (2005, 2007; Long & Crookes, 1992) strongly advocates for a needs analysis based on “the real world target tasks learners are preparing to undertake” (Long & Crookes, 1992, p.44) as the foundation for selecting pedagogic tasks, the teachers did not carry out any formal needs analysis, given that they had to follow the top-down mandated textbooks. They instead, drawing on their strong learner orientation and understanding of what tasks could best engage students, changed or replaced design features of textbook tasks, to induce engaging processes that they believed conducive to maximal output in the target language. In this regard, task authenticity refers to how learners *respond* when doing the task rather than whether the task occurs in the real world or not (Skehan, 2003, p.3).

Taken together, in this EFL context, the teachers’ observed practice and reasoning are most accurately summed up by the notion of ‘authenticity through engagement’, proposed by Guariento and Morley (2001, p.350):

Authenticity of task might be said to depend on whether or not a student is 'engaged by the task'... Ultimately this is probably the most crucial type of authenticity, for unless a learner is somehow engaged by the task, unless they are genuinely interested in its topic and its purpose, and understand its relevance, then the other type of authenticity may count for very little.

The teachers' comments indicate that from their perspective this engagement is social and emotional. While the psycholinguistic rationale behind tasks in task-based language teaching is to foster natural processing of language (Ellis, 2003; Samuda & Bygate, 2008; Skehan, 1998, 2003), the teachers tended to be concerned primarily with how to first engage students socio-affectively as a precondition for language use. This perception is supported by Dörnyei and Kormos (2000, p.281), who, working from learner data, view the role of task engagement in language instruction as "a prerequisite for any language processing to take place". In the current study, the teachers developed practical theories of tasks "while they are on the job" (Kumaravadivelu, 2006a, p.172), grounded in their task experimentation, and an understanding based on experience of their EFL teaching context and learners' interests.

The driving force behind their divergence from the textbook was teacher autonomy, a powerful and legitimate force in the implementation of any mandated approach. As Lamb (2000) comments, "teachers need to understand the constraints upon their practice but, rather than feeling disempowered, they need to empower themselves by finding the spaces and opportunities for manoeuvre" (p.128). The fact that these teachers were doing just that is illustrated by the following comment by one of the teachers in the study:

(10) *Just give teachers a frame, and how this frame is decorated depends on their viewpoints and teaching skills; it's up to them to paint it the way they see relevant; the house is still a house; but this teacher's house is blue, while another's is pink And that's it. (712G)*

This contention captures the reality of classroom teaching in which the teachers always reshape, and re-interpret the top-down textbooks in ways consistent with problem-

solving grounded in pragmatic issues related to their particular teaching context. As another teacher said:

(11) *A task lives its life in class after class,¹² my task implementation is just like an experiment, if the experiment is successful, I will keep it on or otherwise, make necessary modifications or change the task. (411D)*

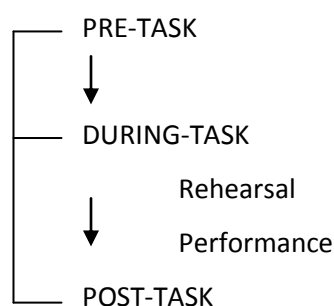
In brief, the study has shown that the nine teachers diverged considerably from the textbook tasks in their teaching from the textbook. In particular, they chose tasks that were more open-ended and more authentic to the students. Underlying these choices made by the teachers was a commitment to find ways to engage the students socio-affectively in the tasks. The socio-affective dimension of tasks has often been overlooked in the TBLT literature and yet it was a fundamental concern to these teachers. The interface between textbook design and teacher practice in this data reflects Prahbu's (1992) insight that "teaching becomes something of an intellectual exploration- a process of subjecting one's theory to an operational test, and sustaining or modifying it in the light of outcomes"(p.239). In a similar vein, Kumaravadivelu (2006a) argues that, "if context-sensitive pedagogic knowledge has to emerge from teachers and their practice of everyday teaching, then they ought to be enabled to theorize from their practice and practice what they theorize" (p.173). The teachers, as we have seen, practised what they theorised and theorised what they practised. Their divergence from the textbook tasks reflects the interaction of theory and practice (Ellis, 1997, 2009b; Erait, 1994) where the former provides 'provisional specifications' (Ellis, 2003, 2009c, 2012) or "can feed *only indirectly* into the practical knowledge that informs actual acts of teaching" (Ellis, 2009c, pp.142, italics added). It is the teachers' pragmatic understanding of how to engage learners that counted directly in their actual practice.

¹² In this high school, one teacher was typically in charge of many classes of the same grade level.

5.4 What task implementation procedures did the teachers use, and why did they use them?

Task design features cannot be fully understood when divorced from the task conditions in which the tasks are carried out (Ellis, 2003, 2009b; Skehan, 1998, 2007, 2009; Skehan et al., 2012). The study also observed the ways the teachers implemented tasks in the classroom. It was found that the teachers organised their task-based lessons into four distinct phases. The first and the last were expectedly pre-task and post-task and in between were task *rehearsal* and *performance*. In task rehearsal students were given some time to prepare and rehearse for the task before they went up to the front of the class to perform it (henceforth, performance). The teachers' task implementation procedures are summarised in Figure 5.2.

FIGURE 5.2: Task implementation procedures by the teachers



5.4.1 Pre-task¹³

In addition to the main oral tasks in each unit, other brief oral tasks were present as warm-up or follow-up tasks for reading, listening and writing lessons. The teachers only prepared students minimally for these tasks and so here only the data from the lessons based on the main speaking tasks are reported. There was considerable variation in the way the nine teachers prepared students for these tasks.

This section covers three areas: the teachers' pre-task work, their underpinning rationales, and students' perceptions.

¹³ This work has been accepted for publication (see Nguyen, Newton & Crabbe, forthcoming). In this work, I collected and analysed the data and wrote the article and my supervisors (Newton and Crabbe) assisted with editing.

5.4.1.1 The teachers' pre-task work

Table 5.4 presents the eight categories of the pre-task work present in the data. It also indicates whether each category was present or absent in the observed lessons by the nine teachers.¹⁴ Data on the time taken for the pre-task phase by each teacher is also provided. The length of the pre-task phase varied tremendously across the lessons, from four minutes in lesson 611F to 24 minutes in 912I. Not surprisingly, the lessons with the shortest pre-task phases (411D, 611F and 712G) contained the fewest distinct teacher actions while the classes with the longest phases (210B and 912I) contained the most. Within this variety of practices there was however a common core: almost all teachers initiated the pre-task phase by introducing the task topic via a short game (a) followed by a brainstorming activity in which the teacher elicited ideas from the class (b). All the teachers introduced the tasks and made sure students understood what to do (c) and a majority of them provided brief suggestions on how to carry out the task (d). Three teachers also set up pair/group structured communication activities as task preparation (e). It is interesting that this emphasis on task content and ideas rather than on language forms/controlled language practice was the single unifying feature of the data.¹⁵ In contrast, language practice was absent in lessons by three teachers (411D, 611F and 712G) and, where present in the lessons by the other six teachers, was approached in a variety of ways. Five of these six teachers introduced some language through teacher talk (f) although only two provided subsequent controlled language practice (g). Four of the five who provided language for the task also followed up by modelling a sample dialogue with a student (h).

¹⁴ The table does not strictly reflect sequence. The actual sequence for all lessons was as follows: a-b-f-g-e-c-d-h, with the only difference being omission of various actions by different teachers.

¹⁵ It is important to note here that although the teachers all taught from the same textbook, they did not plan the lessons together and nor were they following any kind of written guide that stipulated this kind of approach.

TABLE 5.4: Teacher action in the pre-task phase

Teacher action	Teacher								
	110A	210B	310C	411D	511E	611 F	712G	812H	912I
<i>Pre-task focus on meaning and performance</i>									
a. Teacher introduces the task topic with a game	✓	✓	✓	--	✓	✓	✓	✓	✓
b. Teacher leads brainstorming of ideas for the main task	✓	✓	✓	✓	✓	✓	✓	✓	✓
c. Teacher gives task instructions and checks student understanding	✓	✓	✓	✓	✓	✓	✓	✓	✓
d. Teacher briefly provides suggestions for how to do the task	✓	✓	✓	--	✓	--	✓	✓	✓
e. Teacher sets up pair/group structured communication activities to prepare for the main task	--	✓	--	--	--	--	--	✓	✓
<i>Pre-task focus on language practice</i>									
f. Teacher introduces language for the task	✓	✓	✓	--	--	--	--	✓	✓
g. Teacher provides controlled language practice	--	✓	--	--	--	--	--	--	✓
h. Teacher models a sample dialogue with a student	✓	✓	✓	--	✓	--	--	--	✓
Time spent (minutes)	19	23	18	5	7	4	16 ^a	20	24

Note. The tick (✓) here indicates that the teacher took this action and the dash (--) that they did not.

^a In this lesson, 10 minutes was spent on a group game, and this inflates the overall time.

5.4.1.2 The teachers' rationales for pre-task work

I now examine teachers' beliefs and explanations for the pedagogic choices they made in the pre-task phase of the observed lessons. I will focus on the language practice actions (f, g & h) since the teachers varied so much on these, and the students also had clearly articulated positions on the value of such activities. Why did some of the teachers avoid any kind of focus on language form while others included it? It became clear in the interviews that this was not a random phenomenon, nor was it a question of time constraints or lack of relevant expertise. Rather, the teachers all had well thought out positions that underpinned their classroom practices.

Some of the teachers avoided a focus on language form. Here are comments from two of these teachers:

(12) *I don't usually provide useful language before a speaking task because it's not necessary. Students' ideas are varied and rich; they will ask me or friends with some language expressions to convey what they want to, if they feel they need to. (411D)*

(13) *Students help each other when they do the task within their group; students learn from each other in terms of linguistic items and general knowledge. So I just let students do the task in their group first, and correct errors if any later. (611F)*

These two teachers show awareness of the need for an open 'pedagogical space' (Samuda, 2007) in task-based learning. The other two teachers who did not introduce language also explicitly stated that the students needed freedom to carry out the task. Excerpt 1 provides one such example. The main tasks involved students talking about their future jobs and then debating about job choices. Prior to the main task, the teacher had students play a miming game (job guesses). He then briefly elicited students' ideas and then quickly noted down each contributed idea onto the board.

Excerpt 1 (Pre-task- Speaking lesson-712G)

- 01 T: What do you think about 'doctor'? For example, a very amusing job, I think so,
that's *my* idea. What do you think of 'nurse? What do you think of 'nurse? Who,
please?
- 02 S1: Boring.
- 03 T: Hah?
- 04 S1: Boring.
- 05 T: Boring. Thank you. How about teacher? M, please.
- 06 S2: Intelligent.
- 07 T: Yeap. How about farmer? Do you want to be a farmer?
- 08 S3: Strong.
- 09 T: Strong. Good. How about businessmen and (business) women? How about
businessmen and (business) women?
- 10 S4: Rich.
- 11 T: Rich. And the last one, politician?
- 12 S5: Challenging.
- 13 T: Huh?
- 14 S5: Challenging.
- 15 T: Challenging. Ok. Don't worry about your opinions. No right or wrong answers, don't
worry. Now good points, what about teachers?
- 16 S6: Take care of children.
- 17 T: Take care of children? What about doctor?
- 18 S2: Help sick people.
- 19 T: Help sick people, good. A nurse?
- 20 S7: Take care of patients.
- 21 T: Take care of patients, and doctors?
[The whole class laugh]
- 22 T: Farmer?
- 23 S8: Erm erm produce foods.
- 24 T: Produce foods. Businessmen and [business] women? Please, raise your hands. H,
please.
- 25 S9: Develop the economy.
- 26 T: Develop the economy.
- 27 S10: Control the country.

28 T: Control the country?

...

29 T: You can use the jobs and the ideas on the blackboard, or you can use the jobs not on the blackboard, any jobs that you want to do in the future. For example, I want to be a designer, I want to be a model, I want to be a singer, I want to be an engineer or an architect. Discuss with your friend what do you feel about that job, good points and bad points of that job. Ok? Clear? Now you have five minutes to talk to the person next to you first.

In this interaction, the teacher did not specify any language to use. Instead, students themselves, by expressing their opinions, produced adjectives describing jobs ('boring', 'intelligent', 'strong', 'rich', 'challenging'-lines 2, 6, 8, 10, 12) and collocations ('take care of children', 'help sick people', 'take care of patients', 'produce foods', 'develop the economy'- lines 16, 18, 20, 23, 25). The teacher also encouraged talk by saying there are no 'right' or 'wrong' answers (line 15) and made it explicit that students could talk about any job that they wanted (line 29). Talking about his pedagogic moves, the teacher commented:

(14) *Guiding should be in a way so that students actively come up with input. That's part of my viewpoint on learner-centeredness! (712G)*

The other teacher similarly stated:

(15) *Teacher guidance towards the task is very important. I often give students some hints to go about the task, but I don't ever force them to follow my suggestions. Before students do the task, I make explicit what flexibility students can have, you can't force students to follow teachers; I let them add what they feel relevant so that they will be more responsible- as long as they use English and talk a lot- this is my teaching experience which has been built up for many years. (611F)*

The flexible and non-prescriptive approach articulated by these teachers echoes the views of the case study teachers in Andon and Eckerth (2009) who sought to give students "the freedom to say what they want and, to some extent, to decide what language they use" (p.294).

In contrast, those teachers who introduced pre-task language (and controlled practice) worked from a different set of beliefs. Excerpt 2 provides an example. The teacher elicited and presented phases for asking for and giving opinions, before putting students into groups to do the task which was to decide on the best seats on a boat trip for a group of people.

Excerpt 2(Pre-task-Speaking lesson-210B)

- 01 T: If you want to ask for one's opinion. What do you say? N, please.
- 02 S1: What about you?
- 03 T: Right. Another way?
- 04 S2: What do you think?
- 05 T: What do you think? Or what do you think of? What do you think of something? Right? Or what's your ...? may be what's your idea, right? If you want to give an opinion, how do you say? H, please.
- 06 S3: I think.
- 07 T: Very good, I think ... Another way? M, please.
- 08 S4: I think we should.
- 09 T: Right. Very good. I think we should ... I think you should. Do you think we shouldn't?
- 10 S1: No.
- 11 T: Yeah, we say I don't think we should. Any other way?
- 12 S5: In my opinion.
- 13 T: Yeah, any other? You can say 'to my mind', or you can say 'as I see it'. If you agree with somebody's opinion, what do you say? How about the others? Come on! N, please.
- 14 S6: I agree with you.
- 15 T: I agree with you. Very good. Any other? L, please.
- 16 S7: I think so, you're right.
- 17 T: Good.

During this activity the teacher wrote the elicited language items on the board and then had students do a sorting activity in which they listed expressions for giving opinions or advice in two separate columns. Her beliefs about the importance of a pre-task focus on language are consistent with her pedagogy:

- (16) *Pre-task language input is very necessary. Without it, students will find it hard to express themselves in English. Without input, it's like in a vast sea, they don't know which way to go and they struggle. (210B)*

The other teachers who provided pre-task language made similar comments.

Continuing the focus on pre-task language input, five of the teachers modelled a sample dialogue orally with one student in the class (4 of them also introduced pre-task language). Explaining this action they said:

- (17) *I think a model facilitates student task performance. From my teaching experience, I have noticed that some weak students adhered to the model, using some of the structures in the model. While good students might have their own ways of doing the task, weaker ones should have something to rely on. (310C)*
- (18) *Directions are better; the core things (a model and some language input) should be provided so that students can expand language use in their own way based on this; for some students they really need those bases. If not properly guided, students are not able to carry out the tasks efficiently. If students don't have (input) directions to go, they won't be able to do the task well. (912I)*

Comments 17-18 show that the teachers emphasised the value of pre-task modelling for weaker students and as the 'core' for students to expand language production. To illustrate, before putting students into groups discussing their future jobs, one teacher (Comment 18) displayed the following sample dialogue on the white board:

Lan: What would you like to be in the future?

Nam: I would like to be (a) **a doctor**.

Lan: Why do you want to be a doctor?

Nam: Working as a doctor would be (b) **a humane and rewarding job**. I would have a chance to (c) **take care of sick people** and **help save people's lives**.

She then asked one student to act out the model with her. She commented:

- (19) *I want students to replace these ideas (in bold in the model above) with their ideas for the jobs given and make a conversation. In this way, students will get to know the core thing. (912I)*

In contrast, one teacher who did not provide modelling made the following comment:

- (20) *If I give a model or a sample of any kind, it's my own way. 45 students have 45 ways to go about the task, so no need to (model). Besides, students usually have better, more interesting ideas to say, I'm sure. (411D)*

Once again, beliefs informed action. Similarly, the three teachers who provided pair/group structured communication activities prior to the main task (e from Table 5.4) were guided by their belief that these activities facilitated student performance of the main task. One of them said:

- (21) *Those earlier activities help facilitate the main activity (task). Students can use the ideas in the initial activities to do the final one, so it is easier for them, than starting the final activity right away. (812H)*

So while all the teachers embraced a task-based methodology, they varied on the crucial issue of whether to focus on language form in the pre-task stage. Furthermore they were all able to articulate a position on this question that was entirely congruent with their classroom practices. This highlights the role of teacher thinking (Borg, 2006, 2009; Pajares, 1992; Richards, 2008; Woods, 1996) in the implementation of task-based instruction in this context. In the words of Kumaravadivelu (2007):

A central issue in the implementation of task-based instruction is when and how to promote a principled focus on form. To a large extent, one's stand on this issue will shape the nature of task design, syllabus construction and instructional strategies. (p.19)

Pre-teaching linguistic items risks prompting students to see the subsequent task as an opportunity to practice the target items and thus compromise the 'taskness' of the task (Ellis, 2003; Willis & Willis, 2007) or lead to 'de-tasking' (Van den Branden et al., 2009). However, form and meaning are on a continuum. As Skehan (1998, p.96) notes, "avoidance of specific structures and engagement of worthwhile meanings, are matters

of degree, rather than being categorical.” One way to address the question of whether a pre-task linguistic focus devalues the task is to investigate the experience of learners. This is the focus of the next section.

5.4.1.3 Students’ perceptions of the learning opportunities offered in the pre-task work

This section reports on student perspectives on the ways the teachers ran the pre-task phase. Like the teachers, the students were able to articulate clear opinions on the value of a pre-task focus on language forms. Indeed, students had strong, internally consistent but contrasting opinions about all three forms of language practice.

The findings show that two thirds (36/54) of the students expressed a preference for doing the task in their own way without any model, language input or structured communication activities. Interestingly, whether the students viewed pre-task language practice as constraining, facilitating or neutral, they tended to hold parallel attitudes towards modelling (and structured communication activities). The students who held the ‘constraining’ view valued creativity in task performance as seen in the comments below (student identity code in brackets).

(22) *With vocabulary or structures provided, psychologically we tend to use these words, or think in ways that can use these words, so our thinking is constrained. Without them, we can think further with diverse creative ideas.(TQB-12H)*

(23) *For me, it is not good to teach vocabulary or structures before I do the task; it should not be too early. It should be done after I have done the task. From my own experience, if I am corrected certain errors after I have used them in my talk, I will remember them better. So, I prefer to do the task freely, and later the teacher can correct the words that students used incorrectly. (LVT-10C)*

Overall, the students were able to articulate well thought through positions on how they preferred to carry out communicative tasks. Contrary to findings that learners might not see much relevance in oral communicative tasks in EFL contexts (McDonough, 2004; Pham, 2007), the students here appeared to be well aware of the value of these tasks

and how to make the best use of them for language learning. In this way, they added more 'learning value' to tasks (Crabbe, 2003, 2007).

With regards to pre-task modelling, the majority of students saw it as constraining and inhibiting:

- (24) *In my opinion, there should not be a model of task performance, because students tend to rely on it, thus affecting their ability to be creative in language use. For the teenagers' age, creativity is quite big; the only problem that we face is limited vocabulary to express all that we want to mean. However, this problem is not hard to solve. We can always ask peers and the teacher. (BDH-12G)*
- (25) *Without a model, students will have to think in more positive and broader directions. A given model inhibits students from thinking further, and this is very likely to lead to moulding, everybody will do the task in the same way. The imagination of each individual is different, so I believe, without a model we will have a variety of talks (performances), and this is motivating and fun. (LBN-11E)*

Students' opinions here contrast with the positive value placed on pre-task language modelling by students in quasi-experimental studies (Kim & McDonough, 2011). This suggests that greater attention be given to students' perceptions in the TBLT literature.

Littlewood (2004, 2007) argues that structured communication activities offer teachers in Asian contexts a way to gradually introduce TBLT. However, such general advice addressed to teachers to solve a teaching problem was not universally supported by the perception of this group of Asian learners of effective learning. Indeed, most students in the current study held an opposing view. For example:

- (26) *Why not give us this final task right away, no need to do those activities; it's too framing, it's like a pre-determined path to follow. It's constraining and boring. (NTLC-12H)*
- (27) *In those early activities, there's nothing to talk about, just matching ideas and these are not the jobs we want to do, how can we talk about it? (NHAV-12I)*

- (28) *I think just let students do the main task and not do this activity. No need to form questions and answers like this before doing the final communicative activity. It's boring. (LMT-10B)*

In contrast to the students discussed above, 12/54 students expressed a need for pre-task modelling and language (and structured communication activities) as in the following comments:

- (29) *I need the basic thing, (like a model of task performance) from which to expand my talk. In this way, I speak better. (LHT-10B)*
- (30) *My vocabulary is limited, so with some words or structures given, I can use them when needed. So it's easier. (LAD-12H)*

A smaller number (6/54) expressed a neutral position:

- (31) *With some vocabulary or a model provided, those who want to be creative, they can at their own wish; and those who cannot create something new nor want to take risks can use what was provided to help them move on with the task. In my case, I always think of new ideas and thus I rarely use the vocabulary or model provided. It is like the teacher gives you a pen, whether you use it or not is up to you; as long as you still write something in the end. (DTT-10A)*
- (32) *Modelling is just a suggestion. We still have the right to be creative and do the task in our own way. The teacher does not force us to follow the model. (QTHG-12I)*

Overall, these comments from students reflect the value they placed on creativity and on the learning space needed to maximise the benefits from task-based learning. What some teachers did and believed was facilitative of student task performance, was in fact viewed as constraining by many students.

Interestingly, the student voice in the current study challenges earlier views of Vietnamese cultural values such as the notion of “classroom as family”, in which students responded, as a whole group to teacher questions (Kramsch & Sullivan, 1996; Sullivan, 2000). In fact, the student voice in this study supported the challenge that Phan (2004) made against the notion that cultural values (e.g., Confucianism) constrain the

implementation of communicative tasks in Vietnamese university contexts. The data also undermines stereotypical views of Asian students as passive receivers of knowledge (e.g., Cortazzi & Jin, 1996; Flowerdew & Miller, 1995; Pennycook, 1998). Indeed, the students, as the findings have shown, voiced their desire to have space to take an active role in seeking knowledge for themselves. Littlewood (2000) in a survey on students from eight Asian countries found that students were not 'passive' as preconceived. Indeed he found that Asian students do want to be active and independent in learning. As Butler (2011) argues, "it is thus potentially misleading to overemphasise the role of traditional cultural values (such as Confucian values) in shaping Asian classroom practices at all grade levels across Asia" (p.40).

In this study it is students voicing what they experienced as inhibiting or facilitating in task-based teaching that alerts us to the ways pre-task pedagogic actions can thwart or enhance their engagement and learning. Clearly how learning opportunities are viewed is "more a matter of how those present affectively interpret them" (Allwright, 2005, p.22). Importantly, all the learners' perspectives reported here should be situated in the context of the rehearsal-performance model they followed. In the performance where students displayed their work to others, it is reasonable that they would prefer freedom to be creative so as to have different performances. The notion of performance once again emerged as crucial in affecting learners' orientations towards learning opportunities through task work.

5.4.1.4 Pre-task work: Concluding remarks

In the pre-task work, all teachers maintained an initial focus on preparing students for the final public performance. The notion of public performance was highly salient in the culture of all these classrooms and all teachers provided preparation for that publicly performed task in the form of generating ideas and preparatory communicative performance (see 5.3.2 and 6.5).

However there was also considerable variation in the realisation of the pre-task phase across the teachers and the classes both in terms of the length of time allocated to it and in the extent to which teachers included a focus on language forms. These varied practices reflect the contrasting beliefs that the teachers held about language learning

and, therefore the types of learning opportunity to be offered by a teacher. These beliefs align with weak and strong versions of task-based teaching (Ellis, 2003).

Just as the teachers had varying views on pre-task work, so did the learners, although in a roughly inverse proportion. The majority of learners expressed an explicit preference for fewer language-focussed activities. Thus, in the life of a task we have a dynamic in the classroom in which different beliefs about the value of pre-task work are operating. It would be reasonable to assume that these beliefs, in the case of the students, would have an influence on the way in which they take up the opportunity and really use the pre-task work to some advantage.

While it would also be reasonable to assume that such divergences would appear in most classrooms, it is probable that the actual nature of the beliefs are influenced by the nature of the school (in this case an elite school) and the motivation and experience of the students and the experience of the teachers.

That the teachers prepared students for task performance more in the speaking lesson proper than in other unit sessions portrays the pedagogical roles of oral tasks in the context of a unit sequence and provides insights which would not have been possible if observations had been made randomly. The unit sequence offers a broader context both 'spatial' and 'temporal' for more and deeper understanding (Gibbons, 2006, p.5).

5.4.2 During-task: Rehearsal-performance

As mentioned earlier, in the observed lessons, after the pre-task work, the teachers put students into pairs/groups to prepare and rehearse for the task (*rehearsal*) before it was performed in front of the class (*performance*). Performance in the current study appeared to be in line with what Willis (1996) calls 'public report' in the task cycle, where students report the task outcomes to the whole class. However, it was, in fact, typically dyad/group performance as distinct from reporting. The teachers talked about their practice and belief in dyad/group performance:

- (33) *One student reporting the group discussion or whole group performance? It doesn't matter, as long as students produce output, as long as they use English, talk and talk a lot, it's good. But I tend to go for group performance more often as it attracts the attention of the class better.(411D)*
- (34) *I often ask all the group members to go up to the front and perform the task together. In this way, each of them has to be responsible. Besides, when the whole group/pair perform, it is more fun and more interesting as the class can see different traits from their friends. In this way, their classmates will be more attentive in watching and listening.(110A)*

Here the teachers used dyad/group performance as a tool to attract the attention of the class. Notably, performance was located in the during-task stage, which differed from public performance as a post-task activity as recommended in Skehan's (1996a) framework. This is also where Willis's framework and Skehan's differ (see Skehan, 2007b for a detailed discussion of this difference). Table 5.5 locates public performance in the current study, in Skehan's framework and in Willis's.

TABLE 5.5: Locating public performance (adapted from Skehan, 2007b, p.61)

	Teachers in the current study	Skehan's (1996a) framework	Willis's (1996) framework
Pre-task	Activating activities	Pre-task planning	Activating tasks
	Introducing useful expressions		Input exposure
	Modelling		
During task	Rehearsal	Task performance	Task
	<i>Public performance</i>		Planning
	(dyad/group)		<i>Public report</i> (group member reporting on task outcomes)
Post-task	Language analysis	<i>Public performance</i>	Language analysis
	Language consolidation	Transcription	Language practice
	Feedback	Task repetition	

Recall that in the preceding sections, the socio-affective dimension was central in the teachers' choice of tasks. Executing tasks in a rehearsal-performance condition was also

found to be integral and linked to the socio-affective principle that the teachers held. Five main themes emerged from the teachers' comments on performance in the classroom. These were *performance as task outcome*, *performance as motivation*, *performance as happy ending*, *performance as training self*, and *performance as forum for language learning*. Each of these themes is elaborated below.

First, the teachers viewed performance as the outcome or product of the task work:

- (35) *The final product of the task is where students performed up there in front of the class. It shows how students have done the task in their own pairs or groups. (511E)*
- (36) *Students need a place to show their product to the rest of the class. Performance is the outcome of the task, which is evaluated by both students and the teacher. (310C)*

Second, the nine teachers made explicit reference to performance as a way of motivating students to engage in tasks:

- (37) *Having students perform the task in front of the class is a way of encouraging students to engage in oral tasks. They want to be good in other people's eyes, so they have to make learning endeavours when they prepare for it. Otherwise, you know, it's not easy, a shared L1 in the classroom, pen-and-paper exams, no speaking in English outside the classroom. (712G)*
- (38) *If students know that they will have to perform the task in front of the class, they will be motivated or at least 'forced' to work to prepare for it. They will be more engaged and more responsible. (812H)*

Here in the teachers' thinking, performance played the pressurizing role of pushing learners to use the target language. This corroborates what Skehan (1996a, 1998; Skehan & Foster, 1997) claims about the potential of a public performance. In the words of Skehan (1996a, p.56) "the knowledge while the task is being done that a task may have to be re-done publicly will cause learners to allocate attention to the goals of restructuring and accuracy where otherwise they would not." The teachers' beliefs were consistent with student perspectives (see Chapter 6) and echoed Willis's (1996)

comment that “without the incentive of the report, the learning process of planning, drafting, and rehearsing would not happen” (p.58).

Thirdly, besides a problem-solving tool, performance also had affective value in the teacher’s thinking, as the ‘happy ending’ of the task-based lesson:

(39) *I always try to manage time for students to perform the task in front of the class; otherwise, I feel like there is something lacking, like the lesson is incomplete. What a delight to have good student performance for the whole class to watch and listen to. (310C)*

(40) *Performance is the climax of the task. The class have a chance to watch, listen and learn. (210B)*

The notion of performance as the zenith of the lesson was also what students expected (see Chapter 6), as the teacher’s comment below shows:

(41) *There were days (though very occasionally) I didn't have time for many students to perform the tasks in front of the class, students asked me why. My students always look forward to performing the oral task in front of everyone. In their perceptions it is part of what they have been in. (511E)*

That the teachers (and students) perceived performance as the happy conclusion of the task work that they should reach shows how the structure of the lessons had become a social event and part of the culture of these classrooms.

Fourth, the teachers also believed in performance as training self. Six out of the nine teachers most often referred to performance as a tool to train students not only to be natural and confident in using English, but also to be a confident person in life. Some examples follow.

(42) *In real life at some points, students also have to speak in public. They have to experience this and feel confident to do this. Performing in front of the class helps students become not only a confident English speaker but also a confident person in life. (611F)*

- (43) *I try to call on different groups so that students have the experience of performing in public. I want students to be confident. I call on even most shy students. Student confidence grows with experience, and they will soon enjoy presenting and will do it better, with time and practice. This will be good not only for language learning but also for many other things in their life. (210B)*

The additional significance of task performance that the teachers felt in honing students' confidence for life reflects the Vietnamese culture where the teacher realises a dual role: a teacher role of the English subject and a moral role (Phan, 2008).

Finally, the teachers also articulated their beliefs in performance as an environment for learning English through error correction and skill improvement:

- (44) *After students have performed the task in front of the class, I always check the class's listening comprehension, and ask if any language errors have been made and elicit correction. If students give correct answers, I usually give them good marks. In this way, they are motivated to listen to their friends' talk, and through performance, students' listening skill is also improved. (110A)*
- (45) *Students improve their speaking and listening skills while working with their group members, and also while they listen to other groups talking in front of the class. They can learn from them, and errors, if any, are brought to the whole class's attention. (912I)*

Obviously performance was the communal place where the teacher recruited the attention of the whole class by checking their listening comprehension and correcting errors. Error correction after performance became a joint task shared by both the teacher and students:

- (46) *With a large class and within time limit, it's hard to correct all group work. Performance is a point of departure where we co-correct errors and consolidate linguistic items used. The whole class can learn from one or two performances. (611F)*

With the values given to the (public) performance, the nine teachers all adopted a reactive role in response to task rehearsal under way. Typically they circulated among

groups, encouraging them to talk or providing assistance as necessary for students to prepare for that performance. At the post-task stage, their classroom practice also reflected their reasons for adopting a rehearsal-performance approach. This is addressed next.

5.4.3 Post-task

Excerpt 3 illustrates the follow-up roles the teacher adopted after a group of students had delivered their performance in front of the class. The task involved students discussing what they were going to do in the next three days off.

Excerpt 3 (Speaking-Post-task-110A)

- 01 T: Now, tell me some information that you heard from your friends? H, please?
- 02 S1: They decided to go to Thanh Tan.
- 03 T: Thanh Tan?
- 04 S1: Thanh Tan mineral hot stream.
- 05 T: Ah, Thanh Tan hot stream, or mineral stream or Thanh Tan resort, you can say. Yeah
[signalling the student to continue]
- 06 S1: They will bring so much money and erm ...
- 07 T: So much money, did they say 'so much money'?
- 08 S1: No, 'so many money'.
- 09 T: Yeah, they said 'so many money'. You have to be careful, they said 'so many money'.
'Money' is an uncountable noun, 'so much money', or 'a lot of money', OK? Very good!
You got a lot of information. Thank you. And N?
- 10 S2: I have got some information, but it seems like H.
- 11 T: Seems like H, OK. L, please. [L is putting his hand up]
- 12 S3: I have heard V said that he wants to bring the ball to play football
- 13 T: OK. Yeah, we can play football there, is that right?
- 14 SS: Yeah.
- 15 T: The first group had a good plan, they decided to bring some food, a ball, and a
blanket. Now... P said something like 'we must go' to that place [looking at the student,
smiling], here we can use 'I agree with you' or 'I can't agree with you more' because if
you say 'we must', it means you have to follow because your friends force you. Is that
right?
- 16 S2: Yeah

17 SS: Yeah

18 T: Very good. Now the second group, please.

Excerpt 3 shows that the teacher now is back at the front stage, adopting the role of a 'language teacher' (Willis & Willis, 2007), checking students' listening comprehension, and eliciting content information from three students (S1, S2, and S3) (lines 01, 09, 11). She also introduces different ways of expressing the same meanings (line 05). The teacher corrects the error '*so many money*' after she has elicited the content information from one student (lines 06-09). She introduces the correct forms '*so much money*' or '*a lot of money*' after she has explained that '*money*' is an uncountable noun (line 09). Here, the teacher is in a language instructor role addressing the language errors students have made in the context of a particular task performance. It is a reactive role, that is, the teacher is responding to what arose from students' task performance, rather than working to a pre-determined formal syllabus. In line 15, after complimenting the group for having come up with a good plan, the teacher further suggests ways of better expressing '*we must go to*' as a response to a suggestion such as '*I agree with you*' and '*I can't agree with you more*'. This kind of post-task practice in response to task performance is in line with the benefits of performance that follows rehearsal, as stated by the teachers.

5.4.4 Rehearsal-Performance: Concluding remarks

The teachers used public performance that followed rehearsal as a problem-solving tool to engage students in task work in their EFL teaching context. Their belief that tasks should engage students socio-affectively governed their choice of tasks and accordingly their ways of implementing them. Performance became a social event in these EFL classrooms. I would like to capture the spirit of performance in the present context with Excerpt 4, a public task performance by a group of Grade 10 students. The task required students to discuss a plan for a picnic to relax after their examinations. The performance lasted around 2 minutes.

Excerpt 4 (Performance-Group 1-10B)

- 01 SS: Hết giờ rồi cô ơi! (*Time's up, Teacher!*)
- 02 T: Which group volunteers? Well, ... X, your group, please.
[The whole group came to the front of the class.]
- 03 S1: Hello, everyone
- 04 SS: Oh, hello
- 05 S1: I feel so tired after the exam
- 06 S2: Ok, I want to relax
- 07 S1: What's that? What should we do?
- 08 S2: Do you want to go some (.) somewhere?
- 09 S1: Oh, it's interesting. Um I like to go to Danang, a near person,
erm... a near place near our city.
- 10 SS: [Uh ha [laugh]
- 11 S1: xxx I can go to the beach in this three days and eat erm ...
some seafood, and I like .. 'chả cá' noodles [laugh]
- 12 SS: [laugh]
- 13 S2: Oh, I also like to go to Vung Tau
- 14 S4: Yes, ok
- 15 S2: It's a beautiful place that seafood is very cheap and erm ...
- 16 S1: Delicious, very delicious. What about you, P? [she sounds like
she has been there]
- 17 S3: I want to go to Nha Trang because I like Nha Trang beach
- 18 S2: Yeah
- 19 S1: and Nha Trang is very beautiful for ... And you?
- 20 S4: I like go to Da Lat because in Da Lat have a site famous such as
Love Valley, erm... erm ...
- 21 S2: [Love valley, oh! [laugh]
- 22 S4: [Xuan Huong Lake
- 23 S1: Yes, the lake is very beautiful, and you?
- 24 S5: Umh I like to... I like ...
- 25 S2: Singapore? [laugh]
- 26 S5: No. Singapore, no. I like to Nha Trang city because I want to
erm ...go to beach, beautiful beach
- 27 S1: I think the beach in Nha Trang is not more beautiful like beach in

- Danang, it's very beautiful
[sounding proud of the beach in Danang]
- 28 SS: [Laugh]
- 29 S1: In another country in the world?
- 30 S3: Ok, all Las Vegas in Los Angeles or something like ... I can see Paris By Night, a famous programme.
- 31 S1: (*Competing, seizing the turn*) I want to go to Korea to see my biggest idols
- 32 SS: [Laugh]
- 33 S2: [Famous singer?
- 34 S3: [Or handsome?
- 35 S1: Yeah
[The whole class laughs]
- 36 S1: And yes, let's we go
- 37 SS: [Yeah. Bye. See you [Saying goodbye to the class]
[Laugh all the way back]
[Hand clapping from the class]
- 38 T: Very good. That's a good idea (.) Let's go or why don't we go
And when will you go to Nha Trang?[pointing at the student who mentioned this] Erm ...erm... next summer?
- 39 S5: Yeah, maybe
- 40 T: May I go with you?
- 41 S5: Yes. Yeah [sounding very happy]
- 42 S3: 'May I go with you' [Repeating the teacher's question with delighted laugh]
- 43 S1: Cả bây đi luôn đi! (*The whole class should go!*) [sounding very excited][The whole class laughs]

All the laughs and the eagerness throughout show that students were naturally and affectively engaged in delivering their performance. The hand clapping from the audience (line 37) further points to the social nature of the public performance. In the follow-up (from line 38), the teacher not only implicitly corrects the student's error ('*Let's go*' or '*why don't we go*', instead of '*let's we go*') but also focuses on what students have

planned to do, asking when one student will go to Nha Trang (line 38) and whether she could join the student (line 40). The students respond with delight (lines 41-43).

This public performance demonstrates the social interplay between the task, the task performers, and the rest of the class. Public performance surfaced as a social event that both the speakers (task performers) and listeners (the rest of the class and the teacher) looked forward to and wanted to belong to. This is what Clark (1996) describes as “common ground” (p.221) shared by joint people in an activity.

5.5 Summary of the chapter

This chapter examined the teachers’ use of textbook tasks in the classrooms. The study found that the teachers frequently adapted and replaced textbook tasks. They preferred open-ended tasks with divergent outcomes rather than closed tasks with single correct or convergent outcomes. This teaching practice was informed by the teachers’ beliefs and task experimentation in their own classrooms. The teachers consistently articulated the belief that tasks should engage students socio-affectively to motivate students to use the target language. It was the teachers’ choice of open-ended tasks that motivated the Phase 2 study to include task type (convergent/divergent) in the study design.

In terms of task implementation procedures, the lessons contained four phases: pre-task, rehearsal, performance and post-task. Task performance that followed rehearsal emerged as a shared social event that the teachers saw as providing task motivation, skill enhancement, error correction, and the ‘happy ending’ of the lesson, and nurturing student self-confidence. The teachers’ varied pre-task practices reflected their various beliefs in preparing students for tasks. Students’ perceptions of the teachers’ pre-task options revealed some mismatches between the teachers’ practice and students’ preferences. Many students expressed wishes for freedom and creativity to use the language in their own way and to build a unique task performance. Performance became central in affecting learners’ behaviours. The teachers’ beliefs in the potential benefits of the rehearsal-performance approach also mirrored their post-task work.

The next chapter describes some preliminary findings on how learners engaged in classroom tasks in a rehearsal-performance condition.

Chapter 6 THE PHASE 1 STUDY: HOW THE STUDENTS ENGAGED IN TASKS

6.1 Introduction

This chapter reports on and discusses the results on how students engaged in the classroom tasks, particularly in the rehearsal and performance stages of task implementation that the teachers used. It addresses four research questions (RQs):

- RQ1. To what extent did the Vietnamese high school students attend to form while rehearsing for the performance of communicative tasks? If so, how?
- RQ2. To what extent and for what purpose did the students use L1 in task rehearsal?
- RQ3. To what extent did the students use items accurately in performance that had been subject to LREs in rehearsal?
- RQ4. How did the students perceive communicative tasks, task rehearsal and performance?

6.2 To what extent did the Vietnamese high school students attend to form while rehearsing for the performance of communicative tasks? If so, how?

The first research question asked whether and how Vietnamese high school students focussed on form in rehearsing communicative tasks. Over the 48 rehearsals in the data set which lasted, on average, 5 minutes each (see 4.6.4.1, Chapter 4), the learners engaged in 308 LREs. On average in each rehearsal students discussed 6.4 LREs. The standard deviation was large ($SD= 4.78$), and LREs ranged from zero to 23 LREs (see Table 6.1). This variation could be explained by the fact that the data were collected from different tasks with differing amounts of rehearsal and by different groups of students.

TABLE 6.1: Occurrences of LREs

	n	sum	Min.	Max.	Mean	SD
Grade 10	16	78	1	11	4.87	2.84
Grade 11	11	128	3	23	11.63	6.15
Grade 12	21	102	0	11	4.85	3.07
Total	48	308	0	23	6.41	4.78

Table 6.1 shows the students discussed LREs in a number of ways.

Table 6.1: The ways learners discussed LREs

LRE type	Occurrences	
	<i>n</i>	%
ROL	164	53
RT	8	3
SS	30	10
SC	32	10
CO	54	18
CR	13	4
MS	7	2
<i>Total</i>	308	100

Note. ROL = Requesting help from other learners; RT = Requesting help from the teacher; SS = Self-searching; SC = Self-correcting; CO= Correcting others; CR= Clarification requests; MS= Mixed strategies.

They requested assistance mainly from each other, 53% of the time or 164 out of 308 LREs rather than from the teacher (3%). This contrasts with findings of Williams (1999) who found that her ESL learners, especially lower proficiency learners requested help mainly from the teacher. In the present study, because the class was typically large, the chance for each group of students to ask and receive the assistance from the teacher was low. This also suggests that the students were autonomous, and this autonomy can also explain their self-regulating behaviour through self-searching and self-correcting, around 10% each and correcting others 18% of the time (54/308 LREs). These results lend support to a study by Poole (2005) who found that his EFL learners actively discussed and resolved language problems among themselves.

There were only 7 LREs that involved mixed strategies (2%), indicating that LREs were usually brief in the context of rehearsal for a forthcoming performance with a time limit of 5 minutes, on average. Students rarely negotiated meaning through clarification requests, only 4% of the time (13/308LREs). The low frequency of meaning negotiation instances might be because the open-ended tasks that the teachers used were not conducive to negotiation of meaning. Tasks used in negotiation of meaning studies are commonly information-gap tasks with closed and convergent outcomes (Eckerth, 2008; Foster, 1998; Foster & Ohta, 2005; Newton, 2013; Slimani-Rolls, 2005). Another

interpretation is that in rehearsal students were deliberately finding English resources to prepare their performance, thus reducing the need to negotiate meaning for comprehensibility. This finding shows that negotiation was just one of the many ways the learners benefited through doing communicative tasks in pairs/groups, and is in line with previous findings on learner-learner interaction in EFL contexts (Donato, 1994; Foster & Ohta, 2005; García Mayo & Pica, 2000a, 2000b; Gutiérrez, 2008; Kuiken & Vedder, 2005). The open-ended tasks used by the teachers in the current study seem to have potential for L2 learning beyond negotiation of meaning. As Foster and Ohta (2005) argued, whatever label the task might have, the value of the task does not necessarily lie in how it leads to negotiation of meaning. Interestingly, besides the infrequent occurrences of clarification requests in this study, nearly half led to *translation into L1* instead of modified output in the target language. While this result generally confirms research in classroom contexts (Foster, 1998; Slimani-Rolls, 2005) which found output modification subsequent to negotiation of meaning was infrequent (cf. Gass et al., 2005), this does not necessarily mean translation that follows a clarification request is of no value. Instead the result implies that learners in EFL classroom contexts might benefit from the interaction among themselves in different ways rather than through negotiation of meaning.

In brief, the results showed that the Vietnamese EFL learners attended to form during rehearsal for their public task performance in a variety of ways such as initiating assistance from peers and teachers, correcting others and self, although they rarely engaged in 'negotiation of meaning'.

6.3 To what extent and for what purpose did the students use L1 in task rehearsal?

The second research question concerns the quantity and purpose of Vietnamese (L1) use in task rehearsal. The results are presented in Tables 6.2 and 6.3. Data for both turns and words are presented because the turns varied considerably in length. Table 6.2 shows that L1 turns constituted around 43% (964 turns) of the total; L2 made up 38% (860 turns) and the mixed language turns made up 19% (422). On average, in each task rehearsal the students produced 20 L1 turns ($M= 20.08$) and 18 L2 turns ($M= 17.91$), and 9 code-switching turns ($M= 8.79$). A *Friedman* test showed a significant difference in the

size of the different types of turns ($\chi^2 = 21.027, p < .001$). A follow-up *Wilcoxon signed ranks* test indicated that the mean of L1 turns and L2 turns was significantly different from mixed language turns ($Z = -5.117, p < .001$ and $Z = -3.958, p < .001$), but L1 and L2 turns did not differ statistically from each other ($Z = -1.085, p = .278$). This indicates that in task rehearsal students produced L1 and L2 turns in similar amounts.

TABLE 6.2: Amounts of L1 and L2 use by turn in task rehearsal

Rehearsal (n=48)	Turn					
	n	%	Min.	Max.	Mean	SD
L1	964	42.92	1	59	20.08	15.19
L2	860	38.29	3	56	17.91	14.04
Both L1&L2	422	18.79	0	25	8.79	5.75
Total	2,246	100.00				

With regards to the amounts of L1/L2 use by word, Table 6.3 shows that in total students produced more L1 words (55.8%) than L2 words (44.2%). On average, in each rehearsal, students used more L1 words than L2 words. A paired-samples *t*-test showed no significant difference between these two means, $t(47) = 1.643, p = .107$. This again shows that the students used Vietnamese L1 in roughly equal amounts to the English target language.

TABLE 6.3: Amounts of L1 and L2 use by word in task rehearsal

Task rehearsal (n=48)	Word					
	n	%	Min.	Max.	Mean	SD
L1	7,890	55.8	25	473	164.37	122.49
L2	6,251	44.2	19	522	130.22	93.13
Total	14,141	100.0				

By and large, the findings indicate that during dialogic rehearsal the students used L1 substantially. The results contrast with findings of other studies on the amount of L1 use during pair and group work. For example, Storch and Aldosari (2010) found a limited amount of L1 use with 15 pairs of EFL Arabic learners (7% for L1 words, and 16% for L1 turns). Other studies also found similar low proportions of L1 use. For example, Swain and Lapkin (2001) found Grade 8 French immersion dyads (22) in their talk in preparation for a written task, used L1 for 29% of the turns in the jigsaw task and 21% in the dictogloss task. However, the considerable amount of L1 use in the current study echoes what Guk and Kellogg (2007) found of their Korean EFL learners. These researchers found students used L1 in 46.93% of the total utterances produced. Alegría de la Colina and García Mayo (2009) also found even higher amounts of L1 use of 55-78% (calculated out of L1/L2 words) depending on the tasks (jigsaw, dictogloss, text reconstruction). Alley (2005) also found students used English L1 predominantly in groupwork, at 71%, though for different mediating functions. It should be noted that in the present study the data on L1/L2 use was gathered in the context of task rehearsal in preparation for the subsequent performance of the same task, which was different from all the studies cited here where there was no rehearsal and only a single task performance (cf. Swain & Lakin, 2001). Another note was that student groups varied greatly in amounts of L1 use and this also found support in previous studies (Storch & Aldosari, 2010; Storch & Wiggleworth, 2003; Swain & Lapkin, 2000).

The considerable amount of L1 in the rehearsal stage in the present study can be explained in several ways. First, the students used L1 extensively because they treated public performance as final, and rehearsal as preparatory. Second, under time pressure (there were time limits for task rehearsal, on average 5 minutes), the students used L1 to sort out their ideas and marshal language resources to express the messages they wanted to convey. The following excerpt (translated into English) between the researcher (R) and students (focus group interview) illustrates these points.

- R: Do you use Vietnamese when you work with each other?
- S1: Yes, a lot.
- R: Why?
- S1: Time for preparation (rehearsal) is often limited, if we use English right

away, it is very time-consuming. So we use Vietnamese first to be quick to prepare ideas and find English words later.

S2: Also my friends might not understand all that I say if I say all in English.

R: Umh huh. Do you think when you use Vietnamese, you'll lose opportunities to speak English?

S1: Not really, because the final thing is to speak up there in front of everyone.

R: When you are up there, you can use only English?

S3: Yes, because we have already prepared for it!

(Student focus group interview- 11D)

Third, the students' use of L1 reflected their familiarity and comfort working with each other. Research into pair talk in EFL contexts (e.g., Storch & Aldorsari, 2010) has shown that when students become comfortable working with each other, they tend to use more L1 in their interaction.

Notably, in line with studies on L1 use in pair/groupwork, the current data showed the students using L1 for a variety of functions. One such function was to discuss and resolve language problems (LREs). For example, students used L1 to explicitly ask for assistance concerning English words/phrases to express their intended meanings, weighing language solutions and giving explanations. Excerpt 1 displays this.

Excerpt 1 (Speaking-11E)

S3: Air chi hèo?(*Air what?*)

S1: Air pollute ... air pollute phải không?(*is it air pollute?*)

S3: Air polluted

S4: Air pollution. Pollution là sự ô nhiễm
(*Pollution is the state of being polluted*)

S2: Polluted là bị ô nhiễm (*Polluted is passive*)

Here S3 uses Vietnamese L1 to ask for the word that collocates with 'air'. S1 provides the answer with uncertainty which is expressed in L1. Solutions are then offered, and explained in metalinguistic terms by S1 and S2 at the end of the episode.

Students also used L1 to generate ideas, scaffold, and self-regulate as illustrated in Excerpt 2.

Excerpt 2 (Speaking lesson-11D)

- 01 S2: Bởi vì khi đau ốm ... khi ill
(*Because when they are sick ... when ill*)
- 02 S1: They old (.) when they old (.) they old (.) they are old chứ! (*should be they are old!*)
... sick or old!
- 03 S2: They are sick (.) or old (.) their children will nuôi dưỡng (*take care of*) will erm
- 04 S1: Take care of
- 05 S2: Take care of
- 06 S1: Them. Nếu có nhiều con thì erm nguồn lao động sẽ nhiều (*If they have many children, they will have a good labour force*) ... if they have many children
- 07 S2: Công nhân là workers (*'Workers' is workers*) ... have workers
- 08 S1: Then their family sẽ có nhiều người làm việc (*will have many workers*)
Gia đình họ (.) (*Their family*) their family will have nhiều (*a lot of*) a lot of

Here S2 starts the meaning in line 01, almost entirely in Vietnamese L1. S1 then expresses the meaning S2 wants to say in English, but with false starts. S1 later reflects on his language use by correcting himself '*they are old*' instead of '*they old*', with the emphatic Vietnamese L1 '*chứ*' directed to himself. This indicates his obvious noticing of the difference between the target-like form and his language production. At this point, the output that S1 produces at 02 set "noticing" in train, triggering mental processes that lead to modified output (Swain & Lapkin, 1995, pp.372-373). In other words, after going through 'hypothesis testing', S1 has extended his analysis beyond semantic processing to syntactic processing (Swain, 1995). Through this process, S1 is using L1 as a useful cognitive tool for accessing L2.

S2 continues his contribution by starting from what S1 has said, and in line 03, in the middle of his meaning making, he uses Vietnamese L1 '*nuôi dưỡng*' to regulate his L2 search. Sensing his interlocutor is having difficulty finding the needed word, S1 offers the correct phase '*take care of*', which S2 uses it in his talk. At line 06, S1 completes S2's utterance by adding the pronoun '*them*' after the verb, and keeps generating content in

Vietnamese first and then starts mapping that Vietnamese meaning to English words. Similarly, in lines 7-8, S2 and S1 respectively use Vietnamese L1 to retrieve English resources to express the message they want to convey. In brief, Excerpt 2 shows Vietnamese L1 being used as a mediating tool (Lantolf & Thorne, 2007), and lends support to other studies that show learners using L1 to mediate L2 use (Algería de la Colina & García Mayo, 2009; Alley, 2005; Anto'n & DiCamilla, 1999, Brooks & Donato, 1994; Guk & Kellogg, 2007; Storch & Aldorsari, 2010; Storch & Wigglesworth, 2003; Swain & Lapkin, 2000).

The learners' verbal reports further confirm the roles of L1 as demonstrated above. Among the many functions of L1, 29/54 students said that they used L1 to prepare ideas or meanings first and then map meanings to English forms:

(1) Thinking in Vietnamese is powerful. I can think of millions of ideas that my limited English cannot express them all. My friends can help translate what I think into English. (HVT-10A)

(2) I speak in Vietnamese first to search for and present ideas, and after that I turn those ideas into English. It's like matching meanings to the English words. (NKHH-11F)

Obviously the students viewed L1 as a bridge that connects a meaning to be conveyed to the L2 forms to convey that meaning. Students further said they used L1 to sustain communication:

(3) For example, in the middle of communication, I don't know to express certain ideas, I don't know what else to do but use Vietnamese to move on; I cannot let ideas flow out in English, only Vietnamese can help. (HDH-12I)

(4) Sometimes we have to stop talking in English to use Vietnamese to give explanations. Sometimes we have very brilliant ideas but can't express them in English, so using Vietnamese enables us to speak out ideas, which we later translate into English, and move on with our communication. (TTHL-10B)

These comments reveal that student use of L1 was also related to their lack of English resources, or their inability to use English straight away.

Taken together, student voices indicated the role of L1 as a mediating tool. The result corroborates the findings of previous research (e.g., Alley, 2005; Brooks-Lewis, 2009; Kim & Petraki, 2009) where students perceived the usefulness of using L1 in learning L2. Cook (2001) asserts that

Bringing the L1 back from exile may lead not only to the improvement of existing teaching methods but also to innovations in methodology. In particular, it *may liberate* the task-based learning approach so that it can foster the students' natural collaborative efforts in the classroom through their L1 as well as their L2. (p.419, italics added

This does not mean that students should be encouraged to use L1 as much as possible, but suggests that we must accept the “fact of life” (Stern, 1992, cited in Cook, 2001, p.408) that “two languages are permanently present” (Cook, 2001, p.418). It then follows that if students are prohibited from using L1 in EFL classrooms, they will be denied opportunities to use a useful tool (see Alegría de la Colina & García Mayo, 2009; Brooks-Lewis, 2009; Storch & Aldosari, 2010; Swain & Lapkin, 2000). Given that L1 use has been reported as one of the deterrents to the implementation of TBLT in EFL Asian contexts (e.g., Carless, 2008; Pham, 2007) and as one of the teachers' ‘fears’ (Alley, 2005), the findings here concerning L1 use as a cognitive tool further highlight the need to rehabilitate L1 use. Importantly, despite the extensive use of L1 in the course of rehearsal, in the task performance itself, only English was used. This seems to show the potential of a rehearsal-performance model in EFL contexts (the Phase 2 study further investigated this potential).

6.4 To what extent did the students use items accurately in performance that had been subject to LREs in rehearsal?

The third research question asked whether students used in performance the targeted language items in LREs in rehearsal. The data shows that they did, although because the data set contained only 15 matched rehearsals and performances, the results need to be treated as indicative. A total of 39 out of 79 the linguistic items attended to in LREs in task rehearsal (49.4%) were used *correctly* in task performance. Examples 1-2 illustrate the presence of this uptake.

Example 1 (Speaking lesson-11E)

Rehearsal	Performance
S1: Hiệu quả là chi? (How to say effectively?)	S2: I think the most important solution is the (.) raise ... raise awareness of the people in erm about overpopulation.
S4: Hiệu quả là effectively. (It's effectively.)	S1: I think we must use effectively new energy such as energy of the wind.
S1: Use effectively the energy (.) use effectively the energy	

Example 2 (Speaking lesson-10A)

Rehearsal	Performance
S1: We participate in activities and erm we will hungry at night	S1: I think we should bring food and drink because we will be hungry or thirsty at night.
S2: We will hungry ... will be hungry!	S2: Oh, I think that's a good idea.

In Example 1, in task rehearsal S1 overtly asks S4 how to say '*effectively*' in English and the answer is provided. Later in task performance, S1 was able to use 'effectively' correctly straight away without assistance. In Example 2, S1 makes an error '*we will hungry*' which is corrected by S2. In the subsequent public task performance, S1 successfully incorporates the feedback '*will be hungry*' in her talk. These examples show how the learners used collaborative rehearsal to recall, construct and refine linguistic knowledge (Swain, 2000, 2006). Although they used L1 considerably in rehearsal, in performance, only English was used.

The preliminary finding on the take-up, in task performance, of language items focussed on in LREs is intriguing and validates the teachers' practice and beliefs about the role of

task performance as catalysing target language use. This initial result is also in line with the students' positive attitudes towards communicative tasks and their commitment to preparing for and delivering their performance, as discussed next.

6.5 How did the students perceive communicative tasks, task rehearsal and performance?

In answer to the fourth research question, three themes from the interviews will be reported here: perceptions of communicative tasks, awareness of learning opportunities and orientation toward rehearsal and performance.

On the subject of communicative tasks, 42 (out of 54 students who were interviewed) stated that they highly valued communicative tasks, as seen in the following comments:

(5) I think learning English is learning to communicate, that is, I should be able to speak in English. The ultimate aim is to speak naturally when I get out there in life, to communicate with people when I study overseas, or work with foreigners (MBT-10B).

(6) Being able to communicate in English is very necessary, especially those days, when our country is opening its door to the world, good commands of English are associated with job and study opportunities (BVL-12H).

These favourable attitudes towards communicative tasks are encouraging. In this EFL context, the students have little need to communicate in English outside the classroom, and examinations are not task-based. These factors have been reported to affect the uptake of CLT and TBLT in classrooms (McDonough, 2004; Pham, 2007; also see Adams & Newton, 2009; Littlewood, 2007; Butler, 2011 for recent reviews). But in this setting, the learners saw the relevance of communicative tasks to their future life with particular regard to future work and study, and this is consistent with the teachers' viewpoints on communicative tasks in Chapter 5. It is also consistent with the result in a Thai EFL university which found that students valued communicative tasks because they are related to their academic needs (McDonough & Chaikitmongkol, 2007). Recent research in other Asian countries such as in Japan (Hood et al., 2009), and in Taiwan (Chung & Huang, 2009; Savignon & Wang, 2003) has shown students holding similar positive attitudes towards communicative tasks.

The students also perceived peer interaction through communicative tasks as opportunities to develop creativity, with 39/54 students making comments on this. Here are two such comments:

(7) Doing communicative tasks is fun; it helps develop your thinking, and your ability to respond in English in the course of communication. (NTB-11D)

(8) I think pair and groupwork provide opportunities to think creatively, as I have to think of good ideas' to contribute. (BVML-12H)

That students in the current study saw the benefit of tasks in developing creative thinking again speaks to the potential of the open-ended tasks that the teachers preferred. This is also true of the students' preference for no pre-task input (e.g., pre-task language or performance modelling) which they said constrained their ability to express creative ideas (see Chapter 5). While in the current literature, creativity has been under-represented as a benefit of doing communicative tasks, future research might explore how students' perspective on creativity through communicative tasks influences their task behaviour and language use (see Albert, 2011).

A second theme that emerged from the student interview was learning opportunities through communicative tasks. A majority of students being interviewed (40/54) were well aware of the merits of doing communicative tasks for their English learning, as seen in the following quotes:

(9) Communicative tasks involve using knowledge of grammar, vocabulary, and pronunciation, everything in them. Through doing communicative tasks in pairs or groups I can also learn new vocabulary and new grammar from my friends, and improve my speaking and listening. (CMA-10A)

(10) For each communicative task that I do with friends in pairs or groups, I can always learn something, especially some new vocabulary, or maybe some new grammar, or at least I review my English knowledge. Some grammar I know very well, but I use them wrong in speaking. So through speaking, my friends can correct mistakes for me. (LTYN-10C)

(11) I think for a communicative task, it is important to take advantage of the opportunity to speak, pronounce words, and learn vocabulary from other friends, because each person has his or her own strengths in vocabulary and pronunciation, so we can learn from each other. And this is the purpose of a communicative task in pairs or groups. During practice before we go there to perform the task, if needs for new words arise and I can always ask the teacher or friends. (BLQ-12G)

Throughout the comments the students recognised communicative tasks as contexts for using and expanding their English knowledge, including particularly vocabulary, grammar, and pronunciation. It is also apparent that the learners acknowledged both peer and teacher assistance in making learning happen. This is in contrast with earlier studies which report that learners largely relied on the teacher as knowledge provider and do not see peers as a reliable language source (Davis, 1997; Mackey, McDonough, Fuji, & Tatsumi, 2001; McDonough, 2004; Williams, 1999). The reported benefits in the current study are, however, in line with research by Storch (2005) in which students stated they improved grammar and vocabulary through collaborative writing. In the present study, students used collaborative rehearsal to problem-solve their expressions of meanings in the target language in order to convey to a class audience (performance) (also see Chapter 8) and this explains why they more likely saw each other as mutual resources for learning.

Despite the generally favourable attitudes towards communicative tasks, and reported learning opportunities, one factor that many students (28/54) voiced concern about was that peers might have different viewpoints on learning opportunities through tasks. They emphasised the importance of peers sharing the same stance and pulling their weight in task-based work:

(12) It requires collective efforts from all group members. If some group members enjoy opportunities to speak, others don't, boredom and de-motivation will follow. The task will just pass by quickly without any learning or peers switch to use of L1 for fun. (LK-12G)

(13) If all classmates should also like to speak, and do the task decently, it will be good. Usually in a group, there are stronger and weaker students; better students will help weaker ones to speak. This happens only if all of us feel the responsibility to work together. (PTL-11E)

The finding is consistent with studies by Watanabe (2008) and Watanabe and Swain (2008) who found that their learners expressed concern over interlocutors' willingness to talk and 'share ideas'.

A third theme from the interview data concerns the value of (public) performance. Some researchers (e.g., Nunan, 1989) argue that a distinction between task design and methodology in TBLT is redundant. However, Ellis (2003, 2009a) and Skehan (1996a, 1998, 2009) take an opposing view that task methodology (pre-task, during-task and post-task) is different from inherent characteristics of a task and can influence learners' task performance. Rehearsal leading to dyad/group task performance seems to provide a driving force for student engagement. In so doing, it offers a way to 'add fire' to the task (Bygate & Samuda, 2009), leading students to talk and work decently. A majority of students interviewed talked about (public) task performance as a core motivation:

(14) Because everyone in the group/pair has to go there (do the performance), so they have to try even though they don't feel like. This turns out to be good, at least it makes people worry, so they prepare. (PVB-10A)

(15) You cannot go there without anything to say. And you want to say good things, too. So I think showing the task in front of the class makes us search more English words, practice more, help each other more. (TVHQ-11D)

(16) I find presenting in front of the class is useful. If we did not have to perform in front of the class, we would make less effort and talk less. We might easily talk about other things in Vietnamese. The pressure makes us practice using English and prepare more. (TMH-12G)

Clearly students constructively orientated themselves to the public performance which they saw as 'pushing' them to collaborate more to use the target English and to reduce off-task talk in Vietnamese L1 or to work properly. Furthermore, students also said they

looked forward to delivering the performance because it enhanced their general communication skills and confidence:

(17) I really like doing conversations in front of the class. It helps build up my confidence and my communication skill. Those who are shy might not like to perform in front of others, but once they get used to it, they will be confident. (PNG-11E)

(18) I often feel regretful if I am not called upon to perform the task in front of the class, because I want to be there. As long as I am there, I get more and more confident. I also get the satisfaction if I do well. (PHT-10A)

Students typically devised the opening and closing for the performance and acted out in their own ways:

Excerpt 3: Speaking lesson-12I

- S1: Xí lên bắt tay đồ rửa nghe chưa. How are you đồ rửa nghe.
(When we are up there, remember to shake hands and the like. Also remember to ask 'how are you', and the like.)
- S2: Hi teacher, hi kids. You look very beautiful today! [Laugh] [Joking]
- S1: Oh, thank you! [Laugh] [Joking]

Excerpt 4: Speaking lesson-11E

- S4: Lên hết luôn à? Lên chi dữ rửa? *(All of us will be up there? Isn't it too many?)*
- S2: Ba bây đứng một phía, tau một phía.
(Three of you stand on one side, me the other side.)
- S1: Bơ tau nói xong tau đứng hót à? *(After I have finished my part, will I just stand silly there?)*
- S3: À đúng rồi, đứng hình tam giác. *(Ah like this, we should stand in the shape of a triangle.)*
[Laugh]
- S2: Sau đây là một cuộc hội thoại. *(Here is a dialogue.)* [Introducing]
- S4: Phỏng vấn đi. *(Let's do an interview.)*
- S1: Ok, nhớ nhìn mặt nhau nghe. *(Ok, remember to look at each other's face.)*

Overall the students' positive perceptions of task performance following rehearsal were in line with the teachers' dedication to bringing the task through to the public

performance, which they described as the ‘happy ending’ of the task-based lesson in their task pedagogy (see Chapter 5).

6.6 Summary of the chapter

This chapter has described some preliminary findings concerning the benefits from a learner perspective of doing tasks in a rehearsal-performance condition that the teachers used in the classrooms. In the current study, in the course of making meanings in task rehearsal, the students discussed language problems encountered. They did this using Vietnamese L1 to a considerable extent, but L1 played an important mediating function. Evidence indicated that there was uptake in task performance of language features focussed on in LREs in rehearsal. The students perceived communicative tasks useful for their future work and study and recognised the opportunities for learning English through tasks. They perceived performance following task rehearsal as beneficially compelling all group members to take a serious attitude to task work in preparation for the performance in English.

The incidence of LREs in task rehearsal and uptake in task performance motivated and provided the underpinnings for the Phase 2 study, in which these findings were investigated in more detail. Phase 1 mainly observed and explored what was going on as the teachers implemented textbook tasks and students engaged in the communicative tasks their teachers used. The data is valuable because it describes tasks in action in classrooms. However, as pointed out earlier, the rehearsal and performance data were from different student groups across and within each grade level (whose proficiency was not known). These groups of students carried out different tasks in different lessons and with differing rehearsal time limits. Furthermore the data on uptake was limited. Phase 2 therefore extended the Phase 1 study by looking at the effects of tasks (convergent and divergent tasks), stemming from the teacher choice of tasks and student proficiency on uptake in task performance. It also systematically examined the take-up, in task performance, of different types of language solutions (correct, incorrect, and unsolved) the learners arrived at during rehearsal.

The next chapter will address the methodology of the Phase 2 study.

7.1 Introduction

The focus of the Phase 2 study was to examine in detail the extent to which learners carried language items targeted in LREs in task rehearsal through to the public performance. The study particularly sought to investigate how task type and proficiency affected uptake in performance.

In this chapter, I describe the research questions, the student participants, the tasks, the study design, the data collection procedures, and the data analysis. Finally, I discuss the validity and reliability of the research before concluding with a summary.

7.2 Research questions

The Phase 2 study sought to answer the following research questions (RQ):

Task rehearsal:

- RQ1. Did tasks and proficiency impact on the occurrences of LREs in task rehearsal?
- RQ2. How did Vietnamese EFL learners resolve their LREs in rehearsal?
- RQ3. Did tasks and proficiency impact on how learners resolved correctly, incorrectly or left unsolved LREs in rehearsal?
- RQ4. Did the linguistic focus of LREs affect how learners resolved them in rehearsal?

Task performance:

- RQ1. To what extent did learning opportunities operationalised as LREs in rehearsal lead to uptake in performance? Did task type and proficiency impact on the level of uptake?
- RQ2. Did task type and proficiency affect different types of uptake in task performance?
- RQ3. Was there a relationship between types of LRE resolutions in rehearsal and uptake in performance?
- RQ4. Did uptake in task performance differ by linguistic focus?

7.3 Student participants

The student participants in the current study were all Grade 11 non-English-major students from a leading high school in Vietnam, the same high school where data collection for Phase 1 took place. In Vietnam, students of the same grade level are typically at the same age, in this case, 16. At the time of data collection, these Grade 11 students had been learning English as a compulsory subject at school for 5 years (since Grade 6). They all followed the same English textbook and the curriculum as specified by the Vietnamese Ministry of Education and Training.

There were two main reasons for the choice of Grade 11 students. First, the study was conducted at the beginning of the first semester-2011, so Grade 10 students were still new to the school and might not know each other well. Second, Grade 10 students came from different junior secondary schools in the whole province and therefore it was more difficult to access their previous English academic results and their previous teachers. Grade 12 students were not chosen because they were pre-occupied with their upcoming high-stakes graduation and university entrance examinations.

7.3.1 Proficiency

Proficiency has been one of the important factors identified by teachers as affecting their implementation of TBLT in real life classrooms (Barnard & Nguyen, 2010; Carless, 2004, 2007; Le & Barnard, 2009; McDonough, 2004; McDonough & Chaikitmongkol, 2007; Pham, 2007; see Adams & Newton, 2009; Butler, 2011; Littlewood, 2007 for reviews). However, task-based research that addresses this variable is still limited, especially in EFL contexts (see Philp & Tognini, 2009). The current study seeks to fill this gap by investigating the performance of high proficiency dyads (HH), low proficiency dyads (LL) and mixed proficiency dyads (HL).

Proficiency has been operationalised differently across studies (see Thomas, 2006 for a meta-analysis of proficiency evaluation in ESL research). Particularly regarding LRE studies, the proficiency level of participants has been identified by course or class levels (Nassaji, 2010; Williams, 1999, 2001), and by student higher or lower proficiency in relation to each other in the same class (Kim & McDonough, 2008; Leiser, 2004;

Watanabe & Swain, 2007). Following Leeson (2004), the current study used students of the same classes and the same grade level who were of higher or lower proficiency relative to each other. Leeson used the teacher's ratings to pair students, as he argued, it "realistically reflects how L2 instructors may organize learner pairs according to proficiency" (Leeson, 2004, p.75). Yet he posits, "future research could use a variety of more 'objective' or standardized proficiency measures, in addition to instructors' ratings, to perhaps obtain a more 'concrete' evaluation of learners' proficiency" (Leeson, 2004, p.75). In the current study, the learners' scores from school English assessments were used in combination with the teachers' evaluation. The procedures of selecting students were described as follows.

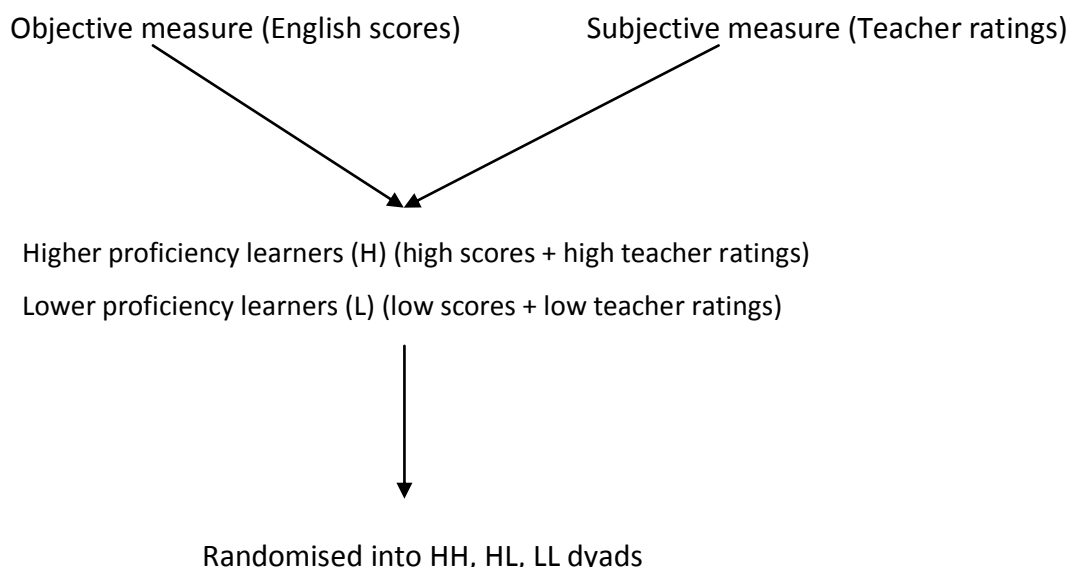
First, volunteer students in six intact classes of non-English majors were selected based on their scores in English¹⁶ in the second semester of the previous academic year.¹⁷ The higher proficiency learners were students who scored from 9.0-10.0 for their English performance and their lower proficiency learners from 6.5 to 7.9.

These overall academic results did not focus on the oral language ability only, but instead covered a range of language areas, namely grammar, vocabulary, reading comprehension, oral performance in the classroom, and regular classroom activity performance. Therefore, the teachers were further asked to rate the students (who were selected based on the overall English academic result), as having higher or lower proficiency in relation to each other in terms of their oral ability. As a result, the students who had both a high English score and high teacher rating were designated as high proficiency. The students with both low scores and low teacher ratings were designated as low proficiency. Then, the finalised high and low proficiency learners in each class were randomly put into three groups (HH, HL, LL). Figure 6.1 summarises the procedures of selecting proficiency groups.

¹⁶ High school students in Vietnam are formally assessed and given the academic English results twice per academic year, divided between Semester 1 and Semester 2, on a 10 point scale (like other subjects). The categorisation of assessment is i) <5.0 = poor performance, ii) 5.0 -6.4 = average performance; iii) 6.5-7.9 = fair performance; iv) 8.0-8.9 = good performance v) 9.0-10.0 = excellent performance.

¹⁷ High school students in Vietnam have around three months' holidays before the new academic year starts.

FIGURE 7.1: Procedures for selecting proficiency groups



Although the number of targeted dyads also varied from class to class, the total number of dyads (8)¹⁸ was the same for each group (HH, HL, LL) (see Table 7.1), thus assuring equal samples, which facilitated greater reliability of statistical comparisons between proficiency groups (Field, 2005; Larson-Hall, 2010). Students were not informed of their interlocutor's proficiency although it was highly likely they were aware, in general terms, of each other's proficiency.

7.3.2 Gender

Although little is known about how gender affects task-based interaction or performance by Vietnamese EFL learners, research in this area shows that gender affects L2 interaction (Gass & Varonis, 1986; Kasanga, 1996; Oliver, 2002; Pica, Holliday, Lewis, Berducci, & Newman, 1991; Ross-Feldman, 2007). In the current study, it was not possible to control for gender due to the availability of volunteer students in each intact class. However, although there was some variation in the gender distribution of dyad members between classes, the total distribution of gender in each proficiency group was

¹⁸There were more student volunteers than the current figures. However, the data of some targeted dyads were excluded due to student absence on one of the tasks or confusing accents (transcriptions were not reliably possible). For the current analysis, although many dyads were of the same gender, fortunately, their distinctive L1 (Vietnamese) and English accents made it reliably possible to transcribe.

almost equal. For example, the HL group and the LL each had three female-female dyads, two male-male dyads, and three mixed gender dyads. Class 11a and Class 11b happened to have only female students. Therefore, for the HH group in the total data, there were four female-female dyads, two male-male dyads and 2 mixed gender dyads (see Table 7.1).

TABLE 7.1: Proficiency groups and gender

Class	Major ¹⁹	HH		HL		LL		Total
		No of dyads	Gender	No of dyads	Gender	No of dyads	Gender	
11a	General	2	2 FF	1	1FF	1	1FF	4
11b	Literature	2	2 FF	0	<i>n/a</i>	2	2FF	4
11c	Chemistry	1	1MF	2	2MF	1	1MF	4
11d	Physics	0	<i>n/a</i>	2	1FF 1MM	1	1MM	3
11e	History- Geography	2	1MM 1 MF	1	1MF	2	2MF	5
11f	General	1	1MM	2	1MM 1FF	1	1MM	4
Total		8	4FF 2MM 2MF	8	3 FF 2MM 3MF	8	3 FF 2MM 3MF	24

Note. F = female, M = male

7.4 The tasks

7.4.1 Task topic

Back in 1989, Tarone and Yule pointed out that controlling the task topic is “all-too-often missing in our interpretation of results” (p.123). They argued that the task topic should be kept constant so that comparisons between performances by different learners could be made. Indeed, some research showed that the topic of a task influenced task-based interaction (Lange, 2000, cited in Ellis, 2003; Newton, 2013). Newton particularly found that his ESL adult learners had significantly more negotiating questions in tasks on a topic about the zoo than in tasks on a medical topic. Lange also found that learners were

¹⁹ At this high school, there were major and non-major classes. Class 11a and 11f were general classes, that is, they did not specialise in any subject. The remaining four classes were majors of subjects other than English.

more engaged in discussing which prisoner to receive parole than which patient to offer a heart transplant. Despite such an important role of the task topic, only a limited number of task studies took it into account (Alegría de la Colina & García Mayo, 2007; Newton, 2013; Swain & Lapkin, 2001). In this study, the two task types were therefore designed to share one *same* topic, the topic of the unit of the English textbook that they were studying (see 7.5, Table 7.2).

7.4.2 Task type

Task type has been reported in task-based research as having predictable effects on negotiation of meaning or focus on form (e.g., Duff, 1986; Foster, 1998; Gass et al., 2005; Gilabert et al., 2009; Lambert & Engler, 2007; Nakahama et al., 2001; Newton, 2013; Slimani-Rolls, 2005) and on overall task performance (e.g., Bygate, 1999; Foster & Skehan, 1996; Kim, 2009; Skehan & Foster, 1999) (see Chapter 3). While these studies have value for language pedagogy, most of them selected task types for different theoretical motivations rather than strictly practical ones.

The selection of the tasks in the present study was not based on a theoretical motivation, but was driven by the teachers' classroom practice in an EFL high school context in Vietnam. The Phase 1 study found that the nine teachers diverged a great deal from the textbook tasks by replacing them with tasks that were more open-ended. In their thinking, they preferred open over closed tasks, divergent over convergent tasks in order to better engage students (see Chapter 5). The phase 2 study investigated the effects of convergent and divergent tasks on student learning (LREs in rehearsal and uptake in performance), because "how teachers cognize must ultimately be considered in terms of the effects their cognitions have on learning" (Ellis, 2009c, p.141).

In this study two types of tasks were designed: one problem-solving task and one debate task. The former involves students reaching an agreement on the task solutions while in the latter, students argue for their different viewpoints, but against their friend's. According to Duff (1986), the two tasks differed in the task goal orientation: the problem-solving task being convergent and the other divergent. Both tasks were open tasks because they did not require a narrowly defined and fixed solution (Ellis, 2003; Pica

et al., 1993), though the debate task was more open²⁰ as it did not oblige students to choose options from those provided like the problem-solving task.

The present study aimed to see the effects, if any, of convergent and divergent tasks on the occurrences of LREs and LRE resolutions in task rehearsal and uptake of the language points in performance. Importantly, rehearsal and performance, once again were the teachers' ways of implementing tasks to encourage L2 use in this homogeneous L1 context. As Ellis (2003, p.337) argues, "if task-based teaching is to make the shift from theory to practice it will be necessary to go beyond the psycholinguistic rationale ... and to address the contextual factors that ultimately determine what materials and procedures teachers choose".

We have seen from Chapter 5 that it was the teachers' thinking, their practical understanding of their teaching situation that drove their autonomy in seeking textbook task adaptations and replacements and ways of implementation (rehearsal-performance) to engage students in task-based work. Again, as Ellis (2010b) argues, "'theory' as understood by researchers is also different from the 'theory' that informs teachers' actions" (p.185). In this light, the inclusion of the two task types in the present study aimed to explore this theory-practice dichotomy.

The two tasks were designed to fit in with the topic of the current textbook unit that students were studying. Since the teachers commonly adapted or replaced textbook tasks, the use of the different tasks from the textbook tasks was unlikely to attract attention from the students.

The main data tasks shared one same topic 'volunteer work'- the topic of unit 4 in their English textbook (see Appendix 7 for a copy of the textbook unit). The problem-solving task requires students to discuss in pairs, deciding on two out of five charity options to spend their 500 million VND on. Both students had to reach an agreement on the two options and give justifications for their choices. The debate task required the dyads to discuss a controversial statement "Charitable giving should be natural to those who are

²⁰ Communicative tasks can differ in degrees of open-endedness, that is, one task might be more open-ended than another, as 'closed' or 'open' are not dichotomous variables (Ellis, 2003; Willis, 1996).

rich". Each of the dyad members was required to take an opposing view and to think of three main reasons to defend their viewpoints (see Appendix 8 for the two tasks).

7.5 Study design

The aim of the current study was to investigate how learners took up, in task performance, language items targeted in LREs in task rehearsal and to identify any task type and proficiency effects. The study adopted a mixed design where task type (problem-solving and debate) was a within-subject variable and proficiency (HH, HL, LL) was a between-subject variable. Data were collected from the six intact volunteer classes (11a-11f). Table 7.2 provides a summary of the design of the study.

Weeks 1-4 served as familiarisation sessions since task familiarity has been shown to have effects on the learners' task performance (Mackey, Kaganas, & Oliver, 2007; Plough & Gass, 1993). The data for the main analysis were collected in Weeks 6-7 (the shaded part) and in between the familiarisation sessions and the main data collection was a one-week break (Week 5).

TABLE 7.2: Design of the study

Time	Grade 11 students					
	Class 11a	Class 11b	Class 11c	Class 11d	Class 11e	Class 11f
Week 1	<i>Topic: Friendship (Unit1)</i>					
	Debate	Problem-solving	Debate	Problem-solving	Debate	Problem-solving
Week 2	<i>Topic: Friendship (Unit1)</i>					
	Problem-solving	Debate	Problem-solving	Debate	Problem-solving	Debate
Week 3	<i>Topic: Party (Unit3)</i>					
	Debate	Problem-solving	Debate	Problem-solving	Debate	Problem-solving
Week 4	<i>Topic: Party (Unit3)</i>					
	Problem-solving	Debate	Problem-solving	Debate	Problem-solving	Debate
Week 5	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>	<i>Break</i>
Week 6	<i>Topic: Volunteer work (Unit 4)</i>					
	Debate	Problem-solving	Debate	Problem-solving	Debate	Problem-solving
Week 7	<i>Topic: Volunteer work (Unit 4)</i>					
	Problem-solving	Debate	Problem-solving	Debate	Problem-solving	Debate

The purpose of the practice sessions was to familiarise students with the task types, task implementation procedures, the recording equipment and the presence of the researcher as non-participant observer before they carried out the tasks for the main analysis. As shown in Table 7.2, students in each class carried out one problem-solving task and one debate task for each of the two textbook unit topics (Friendship and Party) (see Appendix 7) in their normally scheduled class hours in the first four weeks. For the Friendship topic, the problem-solving task involves learners discussing in pairs, deciding on two of the five qualities of a good friend; in the debate task, dyad members, presented their opposing viewpoints on the statement “It’s personality that counts, not appearance”. For the Party topic, in the problem-solving task, dyads had to reach an agreement on two out of the five activities for celebrating a birthday party for one of their friends. The debate task required dyads to be for and against the saying “We must celebrate our birthday because it is a significant event in life”. During these practice sessions, targeted dyads were recorded (with a small digital recorder put on their table) and followed the same procedures as they would in the tasks for the main data collection (see 7.6.5). The purpose was to familiarise them with the recording equipment and task implementation procedures. I, as the researcher, sat quietly in a corner at the back of the class, to observe how students did the task, note the task procedures that the teachers carried out and made sure things did not go wrong technically. It was observed that students did not seem to pay attention to my presence in the classroom.

In order to reduce the practice effect, the order of tasks were counterbalanced between the classes (not proficiency groups), from 11a-11f as shown in Table 7.2, given that it was impossible to do counterbalancing for each proficiency group, since the study occurred in intact classes. As seen from Table 7.2, this counterbalancing was observed right from the familiarisation sessions throughout to the main data. Although the number of dyads for each group slightly varied from class to class, as shown in Table 7.1, in each group (HH, HL, LL) for the total data, there was always half (4 dyads) carrying out the debate task before the problem-solving task and the other half (4 dyads) the problem-solving task before the debate task. In other words, the combined

data for each dyad group were composed of the rehearsals of the problem-solving task or the debate task that preceded and followed the other task type.

In addition, the tasks were carried out one week apart, and there was a one-week break before the main data were collected. The one-week interval between the two tasks thus further helped reduce the possible effect of familiarity with the prior performance on the subsequent performance of the other task.

7.6 Data collection procedures

7.6.1 Ethics

I obtained the approval of the Human Ethics Committee from Victoria University of Wellington for my research before my data collection began in Vietnam (Appendices 1 & 6).

I came to the high school one week before the data collection began. First I sought permission from the Head of school (see Appendix 6.1), and then met with the Grade 11 teachers of English and introduced my research and sought their participation on a voluntary basis. After listening to the verbal introduction of my research and reading the information sheets, three Grade 11²¹ teachers were willing to participate and signed a consent form (see Appendix 6.2). Then I came into their respective classes and informed students of the purpose of the research and asked if they would volunteer to participate (see Appendix 6.3). Six classes (two per teacher²²) were willing. The classes were informed that they would be trying communicative tasks in their regular classrooms. The teacher helped recruit student dyads to be recorded and they told students they would like to see how students of different levels tried out different tasks. Many dyads were willing to participate and signed consent forms. The teachers then noted down the student volunteers in their notebook. They later selected higher

²¹ These three teachers also participated in Phase 1 of the research.

²² In this high school at the time of data collection, one teacher was typically assigned to teaching many classes of the same grade level (also see Footnote 12).

and lower proficiency learners following the criteria described in 7.3.1. Some classes had more dyads of one category (HH, HL, and LL) than the others.

7.6.2 Briefings with the teachers

The teachers were briefed carefully about the task implementation procedures before they carried out the tasks in their classrooms. Classroom observations and audio recordings showed that these procedures were consistently observed throughout the six classes and the two tasks.

7.6.3 The practice sessions

The practice sessions took place prior to the main data collection (see 7.5, Table 7.2). Besides the familiarisation purpose, the practice sessions served as further piloting, to provide feedback on the final design of the tasks for the main data. In the earlier practice sessions, for the debate task, students were left to come up with as many ideas to argue against or for to defend their viewpoints. However, their practice data showed that some dyads went on with a lengthy debate with many arguments, while some came up with just one or two. They explained that this was because the task did not require a specific number of arguments. After the trials in these classes, the task instructions were amended to require each student of the dyad to think of and use three main ideas to argue against each other. It was found that this statement provided better directions for students to follow. The earlier practice sessions also suggested that the data should be collected during a double period, instead of a single one. In this high school, students were scheduled three 45-minute periods of English per week, which are broken into two sessions: one single-period session of 45 minutes, and one double session of 90 minutes, plus a 5-minute break in between. These practice sessions showed that while most classes were able to finish within one single section, some exceeded it by several minutes, as the teacher called upon other pairs besides those being recorded for performance to maintain the normal class life. The main data collection therefore occurred in double sessions of the week.

With regards to the time limit for task rehearsal, after trials and feedback from both the teachers and students, in combination with preliminary analysis of the pilot data, it

was decided that 15 minutes was a reasonable amount of time. One of the students said, “15 minutes is not too much, but not too little. It is just right”. In fact, 15 minutes was not so generous, as the tasks were quite demanding. A 15-minute time limit was kept constant for all dyads and classes, as some studies have shown that time spent on task was positively related to the number of LREs generated (e.g., Storch, 2002b; Storch & Wigglesworth, 2007; Swain & Lapkin, 1998).

In the first two practice sessions, only 5 minutes were allowed for all performances to limit the effect of on-line planning (when no time limit is allowed) (Yuan & Ellis, 2003). Nonetheless, the public performance in the current study encouraged an orientation towards fluency and smoothness. The teachers and students agreed that it was not necessary to specify a time limit for performance. One of the students said: “let us perform what we have, as this is a public performance, we will naturally perform as quickly as we can. We can’t stand there and mutter or search for words to say”. Another said, “when we perform in front of the class, we might forget what we have prepared. And in this case, we use whatever we can to keep the conversation going. We don’t want any breakdown, you know.” Similarly, one of the teachers said, “it is demotivating to stop students if their performance goes beyond 5 minutes. So it is not necessary to specify time for performance.” Furthermore, classroom observations showed that students did not seem to pay attention to the time at all, but just delivered what they had and could. Based on these inner perspectives, in the next familiarisation sessions and in the main study, time for performance was not mentioned. It was found that all the performances naturally lasted a few minutes. For the main data, the mean time students spent on performance was: for the problem-solving task: $M=3.6$ mins ($SD=1.21$) and for the debate task: $M=3.9$ mins ($SD=1.43$). The Pearson correlation tests found no correlation between the time each dyad spent on the (public) performance and the uptake rate, irrespective of the tasks (the problem-solving task: $r=-.159$, $p=.457$; the debate task: $r=-.044$, $p=.840$).

7.6.4 Piloting

Piloting with the two tasks for the main data collection was done with other Grade 11 students in other classes (who did not participate in the current study) to test the

demands of the tasks and task instructions. Feedback was obtained from the teacher and students and some refinements were made to the task input. The piloting showed that students seemed to engage in the tasks, and the tasks were reasonably challenging with 15 minutes for rehearsal prior to performance.

7.6.5 The main data collection

The main data collection occurred in Week 6 and Week 7 of the data collection period after the familiarisation (see Table 7.2) and followed the similar procedures that had been practised in the preceding familiarisation sessions. The targeted dyads (HH, HL, LL) simultaneously carried out the tasks with their classmates in their regular speaking lessons with their usual teachers, except each targeted dyad was recorded with a small digital audio recorder put on their table.

7.6.5.1 Rehearsal

After the teachers had introduced minimally the task and task instructions, they explained the requirements of the task to students, and made sure students understood what they were going to do by checking students' understanding before they began the task. They also briefly explained some of the words in the task input that students might not know or understand. To maximise understanding of the task requirements, the teachers sometimes translated the task input into Vietnamese. All the three teachers spent, on average, 5-7 minutes on task introduction and explanations, consistently following the same procedures. They were also wearing a small digital recorder around their necks, to provide data for further analysis (e.g., cross-checking the consistency in the task implementation in the classes).

Students in the six classes were allowed the same 15 minutes to rehearse for the performance. They were told that they were allowed to take notes in their rehearsal if they wanted, but the notes were not allowed during the task performance. This was also common practice that students had been used to, too. Also, following their usual practice, the teachers put students into pairs and put them as far away from each other as possible.

During the rehearsal, the teachers circulated around the class as usual. Since the focus of the study was on student-generated LREs and how they themselves resolved language problems they were faced with, the teachers were requested to help students only if they asked for it. Thus the teachers refrained from responding to students' errors, or pre-emptively initiating a focus on form.

7.6.5.2 Performance

After the 15-minute time limit had passed, all the students were required to stop talking. The teachers were very strict in ascertaining that all the students had stopped working on their rehearsal before they were called upon for task performance. It had been an established discipline in these classes that when it was time for performance, all students had to stop talking and face the front and listen to their friends' performance. Furthermore, students were well aware that they were trying new communicative tasks in the classrooms, not participating in a test, so they were disciplined once the teacher announced time for performance. However, it was perhaps possible that learners might appear to be listening attentively, but rehearsing silently in their head 'between the lines'. In an effort to minimise this, for the public performance, the order of the targeted dyads being called for the performance in each class was reversed in two ways: between classes and task types, so that the combined data for each proficiency group (HH, HL, LL) included performances that went first, second, third, and so forth. The targeted dyads carried their recorders along with them when they performed in front of the class. They were acquainted with this practice in the preceding practice sessions. Students were found to behave naturally as if they were not being recorded.

At the post-task stage (after the performance), the teachers were requested to ask some overall listening comprehension questions to maintain the ecological atmosphere of a normal classroom. They were asked to avoid incidentally orienting students to being more accurate or more fluent by correcting learners' errors, or eliciting peer students' comments on the accuracy and fluency of the performance, as they might do in their normal hours.

7.6.5.3 Student interviews

Soon after students had completed the two tasks, I conducted informal interviews with them in the form of focus groups to elicit student perceptions of the tasks, task rehearsal and performance. Due to time constraints, and busy class schedules, some interviews were carried out with the whole class in their normal classroom, where students, targeted and non-targeted students, spoke out the answers as they wished. The purpose of the interviews was not to document the learners' moment-to-moment processing (cf. Gass & Mackey, 2000), but to use the tasks that they had rehearsed and performed as contexts to investigate their thoughts and perceptions. Students were asked about what they perceived of the tasks. They were also asked about what they actually did when they were given time to rehearse the two tasks prior to the performance and why they did so (see Appendix 9 for a sample of student interviews). These learners' interview data were used to complement the quantitative analysis.

7.7 Data analysis

7.7.1 The data set

In total, there were 48 rehearsals and 48 corresponding performances for the current main analysis. The rehearsal transcripts totalled 720 minutes (48x15 mins). The performance transcripts in total lasted 179.2 minutes: 86.4 minutes for the problem-solving task and 92.8 minutes for the debate task. For the former: $M=3.6$ mins ($SD=1.21$) and for the latter task: $M=3.9$ mins ($SD=1.43$). The data also included student interview data on their task perceptions and strategies they used during rehearsal for the performance.

7.7.2 Coding LREs in task rehearsal

I randomly transcribed two thirds of the rehearsals and performances in their entirety, and a paid Vietnamese teacher of English transcribed the remaining third. This teacher and I then carefully cross checked all the transcribed rehearsals and performances for accuracy.

All the rehearsals were coded for LREs, which, according to Swain (1998), are “any part of a dialogue in which students talk about the language they are producing, question their language use, or other- or self-correct” (p.70). This definition of LREs takes into account instances where learners, especially in EFL contexts, may use L1 to discuss the L2 (Kim & McDonough, 2011; McDonough & Sunitham, 2009; Storch & Wigglesworth, 2003; Swain & Lapkin, 2000), which was common in the Vietnamese high school as the research context for the current study.²³ In other words, LREs “arise (whether in the L1 or the L2) when learners encounter problems stemming from their ability to interpret and express meaning *in the L2*” (Leeser, 2004, p.60, original italics). To be specific,

LREs include instances in which learners may (a) question the meaning of a linguistic item; (b) question the correctness of the spelling/pronunciation of a word; (c) question the correctness of a grammatical form; or (d) implicitly or explicitly correct their own or another’s usage of a word, form or structure (see, e.g., Williams, 1999 for descriptions and examples of these kinds of LREs). In addition, LREs may include the use of metalinguistic terminology or the articulation of a rule ... (Leeser, 2004, p.56)

In the current study, LREs included all the instances specified by Leeser, with one variation: (a) was expanded to include predominantly learners asking about the L2 (English) equivalents for the L1 meanings they wanted to make or alternative ways of expressing the same meanings. Episode 1 features an LRE where one learner (S1) asks for the English word ‘*encourage*’ to express her meaning, which S2 provides. In Episode 2, learners are together seeking other alternatives to say ‘*everyday*’.

²³ Phase 1 findings showed students used around 55% L1 (Vietnamese) in the task rehearsal (see Chapter 6).

Episode 1 (The problem-solving task-LL-11f)

S1: khuyến khích là chi?

(how to say encourage?)

S2: huh?

S1: khuyến khích là chi?

(how to say encourage?)

S2: encourage

S1: encourage?

S2: encourage (.) khuyến khích (.) cố động a để họ chăm học *(encourage (.) motivate (them) to study hard)*

Episode 2 (The problem-solving task-HH-11a)

S2: keep them going to school

S1: every day

S2: everyday (.) không cái chữ con- *(no another word con-)*

S1: maintain?

S2: không *(no)* (.) con-

S1: constancy

S2: constancy (.) đừng dùng every day nữa, constancy thôi! *(let's not use every day, just use constancy!)*

S1: constantly chơ! *(constantly!)*

S2: ừ, constantly *(yes, constantly)*

In the present data, LREs involved predominantly single episodes. However, where one linguistic point was discussed multiple times in the rehearsal, it was coded as only one LRE. This also included instances where students initiated an LRE and then dropped it to move on with other communication, and then revisited it later. In other words, the LRE boundary was measured by linguistic items discussed whether in L1, L2 or both. However, if different aspects of one language item such as word meaning, pronunciation, or spelling were discussed, they were coded as separate LREs. For example, Episode 3 features two separate LREs: one addresses the English equivalent 'study costs' and the other is the pronunciation of the word 'study'. Similarly, there are two separate LREs in Episode 4: one is the phrase 'heart disease' and the other is the

spelling aspect of the word 'disease'. However, episodes with multiple LREs (Episodes 3-4) were not common in the current data (9 such episodes in total).

Episode 3 (The problem-solving task- HH-11a)

- S1: Chi phí học hành là chi?
(How to say study costs?)
- S2: Costs (.) learn costs? nói học hành là chi? (how to say 'study?')(.) study /ju/
- S1: study /ʌ/
- S2: study costs /ʌ/

Episode 4 (The problem-solving task-HL-11c)

- S2: That will that will save life for the children who unfortunately unfortunately erm have heart heart
- S1: Heart disease
- S2: Ghi rằng tau không biết.
(I don't know how it is spelt.)
- S1: Rì này. [Noting down on a piece of paper]
(It's spelt like this.)
- S2: Heart disease à? Ghi rằng?
(Heart disease? How to spell it?)
- S1: D-i-s-e-a-s-e. [spelling out in Vietnamese L1]
- S2: Heart disease erm they will have a healthy heart.

It is of note that during the course of task rehearsal, students occasionally made requests of the teacher about the words that they did not know to express their communicative intentions. There were 49 LREs of this type, and they were excluded from analysis because the study focussed on how students resolved language problems among themselves. As the focus was on learner-initiated LREs, episodes where the teacher responded to students' errors, or initiated focus on form were not included for analysis.²⁴ Additionally, because the ultimate aim of the study was to trace the transfer of language items discussed in LREs in rehearsal to the oral (not written) performance, 8 spelling LREs, though coded, were not included in the analysis.

²⁴ In this study, the teachers were requested to respond only when asked, and these episodes were therefore rare.

The researcher and another trained²⁵ Vietnamese EFL teacher independently identified LREs randomly from 50% of the data. The percentage agreement was 91%. All disagreements were resolved through discussion, and several cases where no agreement was possible, the data were excluded from analysis. After the inter-coder reliability had been checked, the researcher then identified LREs in the remaining data. In total there were 648 LREs identified. In this study, the occurrences of LREs *per dyad* were counted.

7.7.3 Coding the linguistic focus of LREs

Researchers have varied in the ways they code lexical and grammatical LREs in LRE studies to date. For example, Williams (1999) categorised learner discussion of even functional items such as prepositions into lexical LREs (also see Leeson, 2004) because she said her learners were more concerned with the meaning of the prepositions rather than the grammatical structures (Storch's personal communication with Williams, cited in Storch, 2001c). In contrast, Storch (2001c) coded LREs that involved prepositions as having grammatical focus (see also García Mayor, 2002) because Storch reasoned that when learners in her study discussed prepositions, they seemed to fall on grammatical rules to find resolutions. Some divergence also exists in the classification of pronunciation and spelling LREs. Some researchers categorise these as lexical LREs (Kim & Mc Donough, 2008; Leeson, 2004; McDonough & Sunitham, 2009; Williams, 1999), and others as mechanics-LREs especially in studies that investigate collaborative writing (Storch, 2001c, 2002a, 2002b) (cf. Swain & Lapkin, 1998). However, studies on teacher-learner interaction put each into distinct pronunciation and spelling categories (e.g., Ellis et al., 2001a, 2001b; Loewen, 2005). The current study used the definitions of lexical and grammatical LREs, adopted by McDonough and Sunitham (2009). This choice was made because their definitions are the most widely used, and like the current study they were developed in research on learner-learner interaction and EFL learners (Thai EFL learners).

²⁵ The research met with the teacher and gave her the data, and the definition of LREs, and explained any questions that the teacher had. When the teacher understood what she was to identify (LREs) in the data, she then proceeded with the job independently.

7.7.3.1 Lexical LREs

In this study, “lexical LREs were defined as LREs in which learners talk about or seek the meaning, spelling, or pronunciation of lexical items” (McDonough & Sunitham, 2009, p.239) (also see Kim & McDonough, 2008; Leaser, 2004; Swain & Lapkin, 1995, 1998; Williams, 1999). In the current data, lexical LREs also included instances where students asked about or discussed (1) the L2 (English) word/phrase for the L1 meanings they wanted to make; (2) the meaning of an L2 word/phrase;²⁶ (3) L2 word choice (also synonyms/antonyms/alternative ways of saying same things); (4) the pronunciation of L2 words (e.g., how to pronounce a word, or articulate certain sounds in a word); and (5) the spelling of L2 words. However, all the spelling LREs were not included in the current analysis as reasoned earlier.

Episode 5 features a lexical LRE where S1 asks about the L2 word ‘*care*’ which he does not seem to know or having difficulty retrieving. S2 then provides a solution which S2 uses to move on with his communication.

Episode 5 (The debate task-HH-11e)

S1: I think that they are children so they don’t (.) *quan tâm là chi?* (*how to say care?*)

S2: Care.

S1: They don’t care about money ... the money is not so important. When you play with them they are happy and that’s enough.

Episode 6 is another lexical LRE where one learner asks about the meaning of an English word, which the other learner provides the answer in L1.²⁷ Episode 7 is another illustration of lexical LREs. Here S1’s use of ‘*where have you been?*’ is questioned by S2, who corrects ‘*where have you gone!*’ and S1 then follows.

²⁶ Although this is infrequent in the current data.

²⁷ See Footnote 26.

Episode 6 (The problem-solving task-HH-11a)

- S2: [singing] Condition là chi?
(*What does condition mean?*)
- S1: Điều kiện. [translation]
- S2: Hey?
- S1: Điều kiện. [translation repeated]

Episode 7 (The debate task-HL- 11f)

- S1: Hiền, where have you been? Bạn vừa mới đi đâu về [self-translation]
- S2: Bạn vừa mới đi đâu về à? (*Where have you been?*) (.) where have you gone!
- S1: Where have you gone?
- S2: Bơ tau nói (*Then I'll say*)

In the lexical LRE illustrated in Episode 8, the students deliberate over whether to use 'disabled people', 'the poor', 'the old' or just 'disadvantaged people'. They finally agreed on the latter resolution.

Episode 8 (The problem-solving task-LL-11e)

- S1: it is (.) luckily for you (.) and you should think about the unluckily people. That's the poor (.) the disabled (.) dis-
- S2: Mình dùng disadvantaged (.) được không?
(*Can we use disadvantaged?*)
- S1: Disabled people (.) the poor
- S2: Không, không (.) không cần (.) the disabled vì cái ni có phải
(*No, no (.) no need to (.) the disabled because this is not*)
- S1: Không thì nói rứa đã (*Just say it for now*).The poor (.) the old (.) the disabled
- S2: Những người bất lợi đi (*Let's use disadvantaged people*) (.) disad (.) disadvan ...
- S1: Disadvantaged chơ chi nữa. Ri hớ?
(*For sure, disadvantaged. It's spelt like this?*)
- S2: Ờ (*Yes*).
- S1: Thôi disabled được (*Hey, just disabled is fine*).
- S2: Vì những người bất lợi mà!
(*These people are all disadvantaged people!*)

- S1: Đây là người bất lợi à?
(These are disadvantaged people?)
- S2: Vì toàn bộ những người ni là người bất lợi hết.
(Because these are all disadvantaged people.)
- S1: Ờ (Yeah). Ok (.) I agree with you.

In Episode 9, the students collaboratively discuss how to pronounce the word 'charity'. This pronunciation LRE was also classified as a lexical LRE.

Episode 9 (The problem-solving task-HH-11e)

- S1: Hùng, ah do you think charity là /kæ/ hay là /tʃæ/ ri hê? *(charity is pronounced /kæ/ or /tʃæ/?)*
- S2: /kæ/
- S1: /kærə-/
- S2: [/tʃærəti/ /tʃærəti/
- S1: [/tʃærəti/ /tʃærəti/

7.7.3.2 Grammatical LREs

In this study, following McDonough and Sunitham (2009) grammatical LREs were defined as LREs where dyads discussed aspects of L2 morphology or syntax. These involved the use of tenses, verb forms, verb-subject agreement, passives, word order, comparatives, pronouns, determiners or other function words (see also García Mayor, 2002). Episode 10 is one such grammatical LRE where S1 questions her language use, whether 'they are poor' or 'they poor' is appropriate. S2 provides an answer 'they are poor', with a meta-linguistic explanation that poor is an adjective, not a verb.

Episode 10 (The problem-solving task -HL-11d)

- S1: Ê, they are poor hay they poor thôi hê?
(Hey, they are poor or just they poor?)
- S2: Er they are poor. Poor nó tính từ mà, phải có động từ! *(Poor is an adjective, it needs a verb!)*
- S1: They are poor. They are poor.

Episode 11 describes another grammatical LRE. Here S2 corrects S1 with the pronoun 'THEY' instead of 'it' that S1 uses, as 'they' refers to 'the rich' as a plural noun.

Episode 11 (The debate task-HL-11c)

- S1: the rich when they have a lot of money they will think about themselves. First, they will care about their appearance, their life such as (.) such as buy beautiful clothes, xxx clothes travel around the world (.), taste the most delicious food
- S2: Đến lượt tau! (*It's my turn!*) Oh it's very selfish.
- S1: THEY are very selfish!
- S2: They are very selfish. I think we should à? (*we should?*)

There were some less obvious cases where it might seem that learners accessed these LREs as lexical chunks. Following Storch (2001c), judgement was made as to whether students were resorting to rules to find the solutions or not. If it was more likely that students resorted to L2 rules to resolve a problem, this problem was coded as having a grammatical focus. For example, in Episode 12, S1 is not sure which verb form goes after 'after'. S2 then supplies the answer straight away with the explanation that it follows 'after'. She seems to mean that it follows 'after' as a preposition and the verb must be therefore in the -ing form.

Episode 12 (The problem-solving task- HH-11a)

- S1: and I sell it. After sell the ring I have 5 million VND, xong bơ Khánh nói (*then Khanh will say*)
- S2: you're lucky! [laugh]
- S1: you're lucky chia cho tôi xí, chia cho tôi ít (*share some with me, share some with me*). [laugh]
- S1: after sell dùng thể chi hè? (*which verb form of 'sell' follows 'after'?*)
- S2: after selling, sau after mà (*because it follows after*)
- S1: after selling I have had ...

Another similar grammatical LRE is shown in Episode 13. Here S1 corrects S2's use of the non-targetlike language 'study more and more better' by suggesting 'study better'

only. It was assumed that S1 was referring to the rules for comparative forms of short and long adjectives/adverbs.

Episode 13 (The problem-solving task-HH-11a)

S1: encourage students to study hard

S2: more and more better (.) coi như là những phần thưởng vì những cố gắng của họ (*as rewards for their efforts*)

S1: encourage them to study better thôi! (*just study better!*)

7.7.4 Coding resolutions of LREs

Following other researchers (Kim & McDonough, 2008; Leeson, 2004; McDonough & Sunitham, 2009; Swain, 1998; Watanabe & Swain, 2007), the resolutions of LREs were coded into three types: (1) correctly resolved, (2) incorrectly resolved and (3) unresolved. If many language solutions were proposed during an LRE, the final solution was counted. As the frequency of LREs differed between proficiency groups, LRE resolutions were calculated as a proportion score defined by Kim and McDonough (2008, p.193) as the “number of LREs in each resolution category divided by the total number of LREs”. Like the frequency of LREs, the LRE resolutions *per dyad* were counted.

7.7.4.1 Correctly solved LREs

A correctly resolved LRE was one where “the problem or question was solved correctly either by one learner’s self-correction, either by one learner answering or correcting the other (other-correction)” (Leeson, 2004, p.65). Episodes 1-13 (except Episode 7) above were examples of correctly solved LREs. Episode 14 is another correctly solved LRE. Here, although many language solutions have been proposed, the final one, ‘*heart disease*’ was counted as the correct resolution.

Episode 14 (The problem-solving task-HL-11f)

- S2: who has broken heart à [laugh]
(*who has broken heart?*)
- S1: điên à (.) có phải thất tình đồ chi mô mà broken heart [laugh]
(*Crazy, this is not lovesick at all, not broken heart*)
- S2: rằng hề? (*so what?*)
- S1: ill heart (.) ill (.) i-l-l
- S2: đau (*sick*)
- S1: không! tim bị bệnh, bệnh về tim là disease
(*no! the heart is ill (.) sickness in the heart (.) it's disease*)
- S2: heart disease à? (*heart disease?*)
- S1: heart disease

7.7.4.2 Incorrectly solved LREs

Incorrectly resolved LREs were coded as LREs “in which the learners incorrectly self-repaired, other-repaired, answered a question, or found a solution” (McDonough & Sunitham, 2009, p.240). In other words, incorrectly solved LREs were those “that were resolved incorrectly by one or both of the learners” (Leeser, 2004, p.66). Episode 7 above is also an incorrectly solved LRE, where one learner corrects the other incorrectly. S1 comes up with ‘*where have you been?*’ which is then questioned by S2 who incorrectly corrects it into ‘*where have you gone*’. In Episode 15, S2 assists S1 to express her meaning by providing an incorrect answer ‘*I feel pleasure ...*’, which S2 then acknowledges.

Episode 15 (The problem-solving task-LL-11d)

- S2: Erm now (.) vừa lòng là rằng?
(*Erm now (.) how to say ‘pleased’?*)
- S1: Uhm ý mi nói rằng?
(*Uhm, what do you want to mean?*)
- S2: Tôi vừa lòng với cuộc sống của tôi rồi.
(*I feel pleased with my life.*)
- S1: I feel pleasure erm with my (.) with my life.
- S2: Ừ. (*Yeah.*)

Episode 16 is another illustration of incorrectly solved LREs, where S2 resolves the problem incorrectly: she talks herself into using '*richer*' instead of '*rich people*'. Here perhaps S2 is hypothesis-testing and generalizing the rule of adding the suffix 'er' to indicate person. S1 does not have any response, but continues with the communication.

Episode 16 (The debate task-LL-11b)

S1: because (.)

S2: because rich people have er (.) have a lot of money er (.)

richer người giàu đừng dùng rich people nữa ... (richer means rich people, no need to use rich people)

S1: What else?

Episode 17 is also an incorrectly solved LRE where one learner proposes '*volunteer organism*' to express his meaning '*volunteer organisation*'. Although this is an incorrect resolution, S1 seems to agree ('OK'). In Episode 18, S1 self-corrects her language use incorrectly to '*they will be died*'. As seen from these transcripts, whether the language solutions were completely wrong (in the case of '*richer*') or partially correct ('*volunteer organism*'), they were coded as incorrectly solved LREs.

Episode 17 (The problem-solving task-HL-11e)

S1: Erm erm erm erm organism (.) *tổ chức* (.) *tổ chức* [] *tình nguyện là chi hệ* (*organisation* (.) *organisation* []) *how to say volunteer organisation*) ... volunteer organism (.)
volunteer (.) volunteer organism

S1: Rồi (OK).

Episode 18 (The debate task-LL-11c)

S1: and if we can't help them (.) and if we can't help them nếu chúng ta không giúp đỡ họ (*if we don't help them*), they will die ah they will be died (.) they will be died.

7.7.4.3 Unsolved LREs

Unsolved LREs were LREs "which neither learner could solve the problem nor knew the answer to the question" (Kim & McDonough, 2008, p.189). Specifically in the present

data, these included LREs where students (1) did not know the answer, (2) were not sure about the answer, (3) could not agree on a correct answer, (4) could not find an answer, or (5) just ignored the question. For example, in Episode 19, both interlocutors are trying to search for the English word which means ‘*contribute*’, but fail to find one, even though they find a simpler Vietnamese L1 synonym.

Episode 19 (The problem-solving task-LL-11f)

- S1: So you should er so you should er ủng hộ tiền là chi hèo?*(er how to say contribute money?)*
- S2: Ủng hộ là chi (.) ủng hộ là chi?
(What is ‘contribute’ (.) what is ‘contribute’)
- S1: Ủng hộ là chi? *(What is ‘contribute’?)*
- S2: Ủng hộ là cho? *(It means give?)*

In Episode 20, S1 is not sure whether to use ‘*many*’ or ‘*much*’ to go with ‘*money*’, but his interlocutor just keeps on communicating, which then draws S2 in too, thus leaving the problem unsolved.

Episode 20 (The problem-solving task-LL-11e)

- S1: With very very (.) many (.) much money (.) a (.) very very (.) much hay many (.) much hay many hèo? *(much or many (.) much or many?)*
- S2: Oh
- S1: What will we do
- S2: What will we do

Episode 21 displays another unsolved LRE where S2 helps S1 to express his meaning in English by proposing a solution ‘*immediately*’ which he himself is not quite sure about. The problem is still not resolved.

Episode 21 (The debate task-LL-11e)

- S2: Er that way is .. tức thời là chi hèo (.) tạm thời là ... (*how to say for the time being (.) temporary is ...*)
- S1: Imme immediately
- S2: That time is (.) is
- S1: Phải không hèo (*Is this right?*)... immediately?

In Episode 21, S1 explicitly seeks the English word ‘*depressed*’, but S2 says he gives up and the language problem is not solved.

Episode 21 (The problem-solving task-HH-11f)

- S1: trầm cảm là răng?
(*How to say ‘depressed’?*)
- S2: từ nó chịu a nà!
(*That word - I give up!*)

7.7.5 Coding uptake in task performance

7.7.5.1 The meaning of uptake in the current study

In second language research, the word ‘uptake’ has been used for different meanings (see Smith, 2005 for a review). Originally, Allwright (1984) referred to uptake as whatever learners report they have learnt from a lesson. A few studies have then used uptake in Allwright’s sense to investigate vocabulary learning (Ellis, 1995; Ellis, Heimbach, Tanaka, & Yamazaki, 1999; Slimani-Rolls, 1989). However, Lyster and Ranta (1997) used the term ‘uptake’ to refer to learners’ incorporation of the teacher’s corrective feedback in their immediate utterances. So far, many studies have broadened Lyster and Ranta’s definition to include not just learners’ response to corrective feedback, but also their response to pre-emptive focus (Ellis et al., 2001b). Uptake also had a different meaning in studies that investigate collaborative writing (Brooks & Swain, 2009; Storch, 2002b; Storch & Wiggleworths, 2010; Swain & Lapkin, 2001; Watanabe & Swain, 2007). In these studies uptake was referred to as instances of ‘transfer of knowledge’ formed through collaborative dialogue where students jointly produced a text to their individual writing. Uptake also has an analogous

meaning in studies that report on the transfer of knowledge from collaborative interaction to individual oral presentations (Donato, 1994; Truong & Storch, 2007). The meaning of uptake used in the current study was similar to these latter studies, except that it referred to transfer of knowledge from collaborative rehearsal to public task performance, in this case, as dyadic performance.

7.7.5.2 The process-product approach

Research has sought to forge links between LREs as ‘occasions for learning’ (Swain & Lapkin, 1998) and language learning or development (e.g., Adams, 2007; Loewen, 2005; Nassaji, 2010; Storch, 2002b; Swain & Lapkin, 1998, 2001). Some studies used pre-tests and post-tests to trace learning of language items learners collaboratively discussed (e.g., Swain, 1998; Swain & Lapkin, 2001). However, these studies acknowledged the challenges in designing such tests: it is difficult to foresee what learners are going to focus on in LREs in their interaction and therefore, the number of test items can be small. For instance, Swain and Lapkin (2001) found that the learners made no improvement from the pre-test to post-test scores. They explained that “relatively few LREs were captured in the core test items, so the language learning we had hoped to measure with the pre-test/post-test design could not be revealed in those items” (p.110). This explanation suggests that pre-test-post-tests might not be effective to trace learning to collaborative work, especially for unfocussed tasks (cf. Mackey, 1999). Because of the unpredictability of LREs that makes pre-tests not possible (Swain, 2001), some studies managed to use individualised post-tests only and they were able to show evidence of learning (e.g., Adams, 2007; Loewen, 2005; Nassaji, 2010). However, these researchers pointed to the mismatch between the nature of the LRE-specific post-tests and the nature of LREs that arise out of efforts to communicate in the L2: the former might tap on the explicit knowledge while the latter implicit knowledge. Addressing this disadvantage, two studies (Loewen, 2007; Williams, 2001) attempted to trace language items focussed on in LREs in subsequent spontaneous use, but found only a low percentage of language items featured in LREs were used in subsequent speech. Loewen (2007) explained that “.... a lack of use of the targeted forms does not necessarily indicate an inability to use those forms; it may

simply be that learners had no occasion to use them” (p.114). Clearly, the challenge of measuring learning from LREs remains.

To address this challenge, the current study used a process-product approach to trace learning. This process-product approach was used in previous work on written tasks (Brooks & Swain, 2009; Storch, 2002b; Storch & Wiggleworths, 2010; Swain & Lapkin, 2001; Watanabe & Swain, 2007) and oral tasks (Donato, 1994; Truong & Storch, 2007). In these studies, the approach involved identifying LREs as opportunities for learning (Swain, 1998, 2001; Swain & Lapkin, 1998) that learners create during their ‘collaborative dialogue’ and subsequently tracing the use of the language items arising from them in subsequent *individual* written performance or oral presentation.

The current study sought evidence of the take-up of learning opportunities created through collaborative task rehearsal in the subsequent public task performance, as *dyadic* (not individual performance). Although numerous studies have used this process-product approach in collaborative writing research as cited, few have done so for oral interaction research (Donato, 1994; Truong & Storch, 2007). Furthermore, these two studies were small-scale, focusing on one group of three students (Donato, 1994), and five groups of 4-5 students (Truong & Storch, 2007) pre-planning one task. The relationship between uptake and how learners resolved their LREs in the preceding interaction were not analysed in these studies. The current study systematically linked uptake in task performance back to how LREs were resolved in task rehearsal. The detailed analysis of uptake is described in the next section.

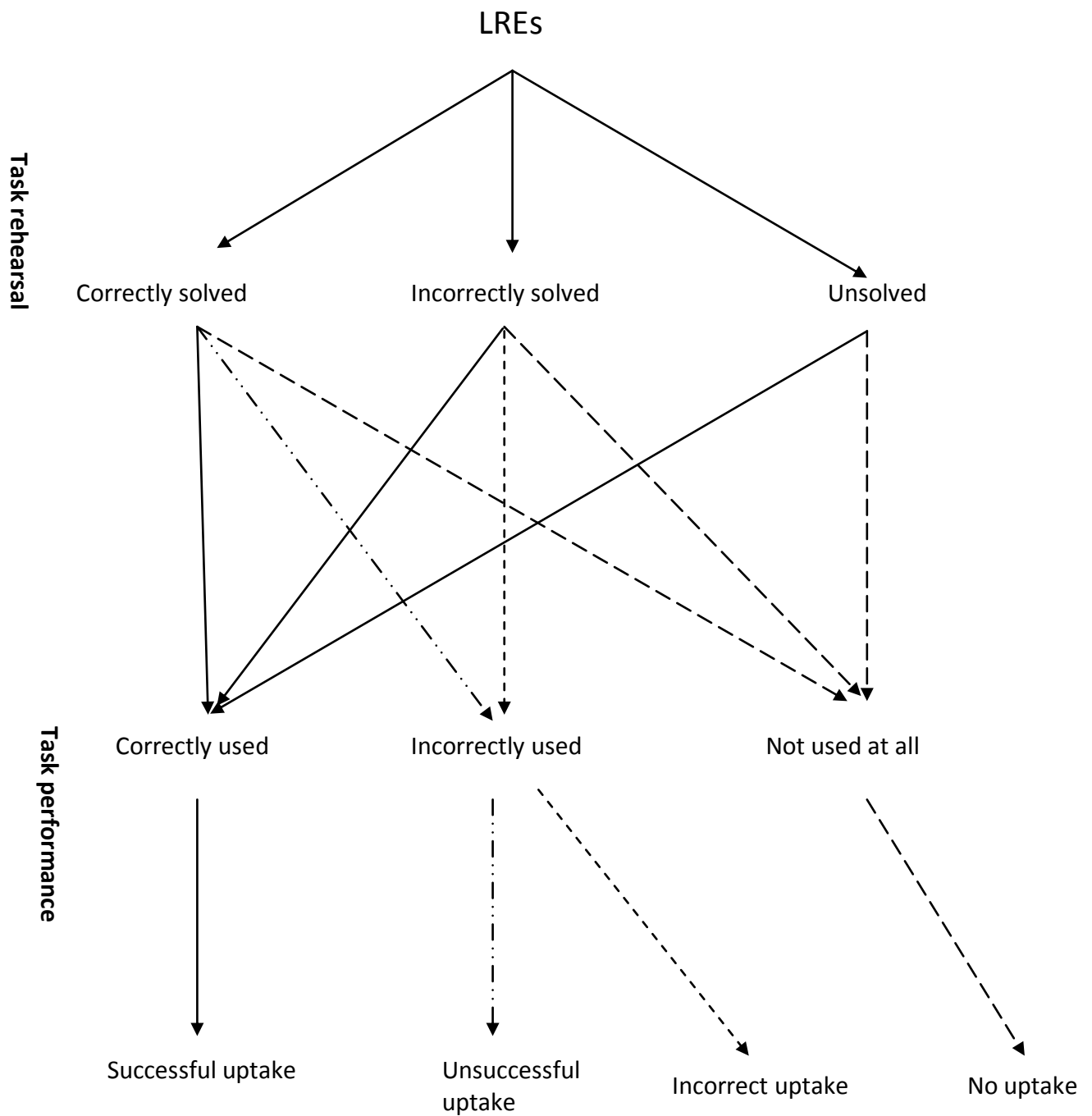
7.7.5.3 Coding uptake

In this study, I used similar procedures of tracing uptake to Storch (2002b) because these procedures presented explicit coding procedures, though focussed on writing. Specifically, LREs were identified as ‘opportunities for learning’ or ‘sources of learning’ (Swain, 1998, 2001; Swain & Lapkin, 1998, 2001) and the language items discussed in such LREs as “potential tracers” (Storch, 2002b, p.314) to trace uptake in task performance. As mentioned above, if a language item was discussed several times during the rehearsal, it was coded as one LRE, and thus one ‘potential tracer’ only (Storch, 2002b).

Where these potential tracers were used in the public performance, they were coded as 'matched items' (Truong & Storch, 2007) or instances suggesting transfer of knowledge (Storch, 2002b) or uptake. It was not necessary that dyads would use all the language items targeted in LREs in their public performance. In other words, only a match of language items that were both focussed on in LREs in the rehearsal and used in the subsequent performance were coded as uptake. It was the pattern in the current data that where uptake occurred, it was for the original interlocutor who encountered the language problems and triggered an LRE. This was regardless of where the solutions were from: from self, the other interlocutor, or co-constructed solutions. Like the LREs and LRE resolutions, the total instances of uptake and no uptake *per dyad* in each proficiency group (HH, HL, or LL) were counted, as the purpose of the study was to make comparisons between dyad groups of different proficiency levels rehearsing and performing two communicative tasks. The evidence of uptake was calculated as proportion scores (i.e., the number of the taken up items out of the total number of LREs).

Coding uptake in task performance was an iterative process of reading, re-reading and identifying the matched items' and labelling them. Figure 7.1 presents the scheme for coding uptake in task performance. Instances of language items targeted in LREs that were not used at all in any way in the performance were coded as no uptake, whether they were correctly solved, incorrectly solved or unsolved in the rehearsal. If the language item featured in a certain LRE in task rehearsal or the 'potential tracer' was used in task performance, it was coded as uptake. In addition to coding the presence or absence of uptake, I also coded its quality in relation to how the 'potential tracers' were resolved in rehearsal and taken up in task performance. In the present data, I coded uptake into three types: successful uptake, unsuccessful uptake and incorrect uptake, each of which is described below.

FIGURE 7.2: The scheme for coding uptake in task performance



Successful uptake

Successful uptake was broadly coded as correct use, in task performance, of language items focussed on in LREs in rehearsal, be they correctly solved, incorrectly solved or unsolved. It was found that none of the unsolved LREs led to uptake. Successful uptake

was therefore coded as correct use of language items that were either correctly solved or incorrectly solved in the rehearsal. The former made up the majority of the successful uptake category. Example 1 features successful uptake.

Example 1 (The problem-solving task-HL-11c)

Rehearsal		Performance	
S2:	Hi Tú ²⁸ [laugh]. Last night did you watch ah(.) the programme?	S2:	Hi Tú
S1:	the auction	S1:	Hi
S2:	huh?	S2:	ahm did you watch er the charity did you watch the charity auction on TV last night?
S1:	auction auction	S1:	yeah I think that's a great programme.
S2:	option? ghi rằng?(<i>how to spell it?</i>)	S2:	after after watching the programme I hope I can do charity, too. Erm if you if you have 500 million VND, what (.) how will you spend it on do doing charity?
S1:	a-u-c-t-i-o-n. auction a!(<i>auction!</i>)		
S2:	Rồi, chi auction chi? (<i>Ok, what auction what?</i>)		
S1:	charity auction		

In Example 1, the language solution ‘*charity auction*’ is correctly solved, and taken up correctly in task performance, and thus the uptake is successful. Similarly, in Example 2, S2 in the rehearsal does not know the word ‘*guilty*’ and seeks assistance from S1, who provides a correct answer ‘*guilty*’ that S2 uses in her immediate utterance. Later in the subsequent public performance, she is able to use it successfully.

²⁸ The real name of the student has been changed.

Example 2 (The debate task-HL-11d)

Rehearsal	Performance
<p>S2: Khi nó tau (<i>Then I'll say</i>) in the past ah they did something wrong and they er (.) regret doing it they feel tội lỗi là chi? (<i>how to say 'guilty'?</i>)</p> <p>S1: Guilty.</p> <p>S2: Guilty. They feel guilty and they want to thoát khỏi ám ảnh (<i>escape from the obsession of</i>)</p>	<p>S2: Er yes, but I think in someone in the past they do they do in the past they did something wrong. Now they feel guilty they want to escape from their fault. Erm they do charity to erm to feel better and leave happiness for their child (.) children Ok. That's a nice idea. Ok. We almost run out of time now. So thank you very much.</p>

Example 3 is another illustration of successful uptake. Here S1 makes a mistake, 'so does' which he is able to self-correct correctly to 'so do' because the subject 'rich people' is plural. He later maintains this correct solution in his performance.

Example 3 (The debate task-HL-11c)

Rehearsal	Performance
<p>S1: I think doing charity is (.) doing charity is (.) is natural for every (.) one. Everyone wants to help people, so does (.) so do the rich. Erm they have even more conditions to help the help people who are living in poverty. Everyone</p> <p>S2: Tau phản đối người giàu à? (<i>Am I against the rich?</i>)</p>	<p>S2: Yes, OK. what do you think about this idea?</p> <p>S1: I think that this idea is er true. In fact you don't need to be rich er to er feel your duty to help people who are living in poverty. Everyone wants to help people, and so so do the rich. They have even more conditions to help the people. And your idea?</p>

As mentioned earlier, if a language problem that was incorrectly solved in the rehearsal got used *correctly* in performance, it was also coded as successful uptake. There were 8 instances of these in the whole data. Example 4 is one such successful uptake. Anticipating S1's difficulty in finding the English word, S2 provides the solution 'terrible', which S1 then modifies 'terribly' (an incorrect resolution), which is repeated

by S2 as a seeming confirmation. However, later in the performance, S1 correctly uses the correct word ‘*terrible*’.

Example 4 (The debate task-LL-11a)

Rehearsal	Performance
S1: <i>everything xảy ra đối với tôi (everything that has occurred to me)ah occur to me is erm erm ter ...</i>	S2: Uyen, why are you so sad?
S2: terrible	S1: I feel in this life everything is terrible.
S1: terrible (.) terribly	
S2: terribly (.) terribly (..)	

Unsuccessful uptake

If a correct language resolution arrived at in task rehearsal got used *incorrectly* in task performance, it was coded as unsuccessful uptake. Example 5 is one such instance where the correct solution ‘*fishing boat*’ in the rehearsal gets used incorrectly in the performance, ‘*fish ... fish boat*’.

Example 5 (The problem-solving task-LL-11e)

Rehearsal	Performance
S1: Yes. I think er (.) some people in a Central Vietnam is er is no is no home and er er they live in a ... <i>đò (boat)</i> ... fish (.)	S1: Oh, it’s perfect I think er some people in Central Vietnam is very poor and and they no no home and they they must live in er a fish ... fish boat
S2: Fishing boat	S2: Yeah. Er and er many typhoon or (.) storm ah usually er get to this land. Ah they must (.) they make the people here (.) more poor
S1: Fishing boat. And er they er (.)	

There were three instances where learners attempted, in their task performance, to use the item targeted in LREs but failed to, as indicated by a long pause. These were

also coded as unsuccessful uptake. Example 6 illustrates this. The rehearsal shows how S1 helps S2 with the phrase *'follow the trend'* which S2 finally understands. However, in the performance, S2 fails to use it, as indexed by a pause which S1 skilfully fills in with a question and S2 naturally responds and their performance goes on.²⁹

Example 6 (The debate task-HL-11b)

Rehearsal	Performance
S2: er but không (<i>no</i>) but they have to (.) không (<i>no</i>). But they think (.) maybe they think they have to do the same (.) với xu hướng như rứa (<i>follow the trend</i>)	S2: they just they do charity just because just because they want to ...
S1: do the same (.) to follow the trend	S1: follow the trend?
S2: to follow (.) to follow the trend	S2: yes. Because they see many people do that and they also do
S1: cụm từ hay rứa mà không nhớ (<i>this is a nice phrase but you didn't you remember</i>)	
S2: ghi rằng? (<i>how to spell it?</i>)	
S1: t-r-e-n-d	
S2: ah xu hướng, rồi, biết rồi! (<i>ah trend, ok, got it!!</i>)	

Incorrect uptake

Incorrect uptake was coded as incorrect use of the language items that were incorrectly solved in the rehearsal. The name 'incorrect uptake' might sound strange, as in this case it was in fact 'successful uptake', though of incorrect resolutions. However, the name 'incorrect uptake' was used in order to distinguish the nature of the types of uptake in the current study (e.g., unsuccessful vs. incorrect uptake). Thus, if an incorrect resolution found in the rehearsal was taken up in the performance in its incorrect form, it was coded as incorrect uptake, as illustrated by Figure 7.1. Examples

²⁹ Despite S1's failure to use the targeted phrase, Example 6 was coded as unsuccessful uptake, because the phrase appeared in the performance (unlike no uptake where the language items were not present at all).

7-8 illustrate this incorrect uptake. In Example 7, in the rehearsal, S2 reflects on the way he uses ‘one three’ to mean 1/3 by way of L1 use. S1 questions it, but S2 responds with an assertion ‘one three’. S1 finally seems to be led by S2. This incorrect ‘one-three’ language solution is later used in the performance by S1. Similarly, in Example 8, the incorrect language solution ‘have lucky’ that came about as S2’s answer provided to S1’s appeal for assistance is maintained in the performance by S1.

Example 7 (The debate task-HL-11f)

Rehearsal	Performance
<p>S2: Thì đó là lý do đầu tiên, có nhiều tiền <i>(That's the first reason, having a lot of money)</i>. Oh my grandfather has do charity every month em [] he erm he use one three tức là một phần ba <i>(that is one-third)</i> his salary to do</p>	<p>S2: Uhm I think only the rich can do charity because they have a lot of money and uhm em I think your family are rich and have a lot of money, don't you?</p>
<p>S1: Chi? <i>(What?)</i></p>	<p>S1: No, for example my family is normal and my grandfather use his salary erm he use one three of it to do charity xxx give to the orphanage or he buy some gift to the poor ... and I don't think that every rich only rich people can do charity.</p>
<p>S2: One-three. S1: One third! S2: One three là một phần ba <i>(One three means one-third)</i></p>	<p>What what do you think ... about it?</p>
<p>S1: Rồi <i>(OK)</i></p>	

Example 8(The debate task-LL-11d)

Rehearsal	Performance
S2: gặp may mắn là rằng? (<i>how to say 'have luck'?</i>)	S2: but if you don't have money how can you do charity work?
S1: have er lucky in life có may mắn trong cuộc sống [self-translating]	S1: charity work is not (.) not only er about money, erm we can help them er in daily activity.
S2: rồi and đồng thời er er moreover là ngoài ra à? (<i>OK and simultaneously er er moreover means besides?</i>)	S2: yes I er I think the rich must do charity work because they ah will have er have lucky lucky in their life and erm they maybe erm erm more and more rich in the future.

In a few cases the learners took up the incorrect form/information, but used it in a different sentence or context. These were also coded as incorrect uptake (Example 9). Here the incorrect information that '*charity*' is a verb is established in the rehearsal and carried to the performance. However, the second time around, S2 uses '*charity*' in a different context '*will charity*', instead of '*how do they charity?*' as in the rehearsal.

Example 9 (The debate task-LL-11b)

Rehearsal	Performance
S1: the poor have (.) don't have enough money to eat	S1: Mai, erm (.) uh money is very important in my life?
S2: họ làm từ thiện bằng cách nào là how do they charity? (<i>'how do they do charity' is 'how do they charity'?</i>)	S2: yes
S1: how DO THEY DO VOLUNTEER [Emphasizing]	S1: If you have erm a large money ah no(.) a large amount of money, what will you do?
S2: charity là cũng làm từ thiện rồi (<i>charity is also a verb already</i>)	S2: uhm yes I will charity
S1: rồi(OK)	

7.7.6 Inter-coder reliability

As mentioned earlier, in total the data were composed of 48 rehearsals followed by 48 corresponding performances. A trained Vietnamese teacher (the second coder) and I coded independently 18 randomly selected rehearsals (associated LREs) and 18 corresponding performances from each of the proficiency groups. These transcripts made up 37.5% (18/48) of the data. The randomly selected data included six rehearsals and six equivalent performances for each of the three proficiency groups (three for the problem-solving task and three for the debate task), totalling 214 LREs. In this manner, the selection of the data for double-coding enhanced 'representativeness' (Révész, 2012). In order to prevent coding towards expectations (Révész, 2012), the second coder was 'blinded' about the proficiency of the learners. After several training sessions, the second coder coded the selected sample (214LREs) into lexical and grammatical LREs, and then coded these LREs into how they were resolved, namely correctly resolved, incorrectly resolved and unresolved. Finally she coded task performance³⁰ for uptake, successful uptake, unsuccessful uptake, incorrect uptake and no uptake, as operationalised above.

The inter-coder reliability scores were calculated using both percentage agreement and Cohen's kappa coefficient. By way of illustration, Table 7.3 features the results of coding LRE resolutions of a task rehearsal which contained 22 LREs by Coder 1 (the researcher) and Coder 2, with 1 = correctly solved, 2= incorrectly solved and 3= unsolved. In the agreement column, if there was no difference between the two coders, 0 was entered to indicate agreement. Where there was a difference between the two coders, in this example, 1, -1, or -2 was entered. The percentage agreement is the sum of 0s divided by the total number of scores available multiplied by 100. In this case the sum of 0s is 19, and the total number of scores available is 22. The percentage agreement is therefore $100 \times 19 / 22 = 86.4\%$. In fact, after the reliability data from the two coders had been made available as demonstrated in Table 7.3, they were submitted to the SPSS 18.0 version for the calculation of both the percentage agreement and the kappa coefficient (κ). For this example, the kappa was .70.

³⁰ The LREs were kept for the rehearsal and the corresponding performance to allow for systematic tracing of uptake.

TABLE 7.3: A sample of reliability data for coding LRE resolutions

LREs The problem- solving task-LL- 11a	Coder 1	Coder 2	Agreement 0=yes 1,-1, -2 =No
1	1	1	0
2	1	1	0
3	1	1	0
4	2	2	0
5	1	1	0
6	2	2	0
7	2	1	1
8	1	1	0
9	1	1	0
10	3	3	0
11	1	1	0
12	1	1	0
13	1	1	0
14	1	1	0
15	1	1	0
16	1	1	0
17	1	1	0
18	2	2	0
19	1	1	0
20	1	2	-1
21	3	3	0
22	1	3	-2

The reliability scores for all the double-coding categories were summarised in Table 7.4. Overall, the percentage agreement ranged from 93.5% to 95.3%, and the kappa values from to .77 to .89. These reliability scores showed satisfactory reliability (Miles & Huberman, 1994;³¹ Viera & Garrett, 2005³²). After assuring the sufficient reliability standards, the researcher coded the remaining data.

³¹ Coders should reach an agreement for a minimum of 80% of the time (Miles & Huberman, 1994).

³² $k < .00$: less than chance agreement; $k = .01$ to $.20$: slight agreement; $k = .21$ to $.40$: fair agreement; $k = .41$ to $.60$: moderate agreement; $k = .61$ to $.80$: substantial agreement; $k = .81$ to $.99$: almost perfect agreement (Viera & Garrett, 2005, p.262).

TABLE 7.4: Reliability results

	Feature	κ	Percentage agreement
Task rehearsal	Linguistic focus of LREs	.77	95.3
	Resolution of LREs	.82	93.5
Task performance	Uptake	.89	93.5

7.7.7 Statistical analysis

Descriptive statistics were obtained to provide information about means, standard deviations and skewness. Repeated Measures Analysis of Variance (RM ANOVA) was the main statistical test used to analyse the data, with one main within-subject variable, task (with two levels: the problem-solving and the debate task), and one between subject-variable, proficiency group (with three levels: HH, HL, LL). All the assumptions of the RM ANOVA (normality, homogeneity of variances, and sphericity) were checked before analysis. The Skewness-z score test of normality (Field, 2005) showed that the assumptions of normality were met for a majority of variables, particularly those with relatively high frequencies (e.g., LREs, lexical LREs, correctly solved LREs, successful uptake). For some variables, the assumptions of normality were not met, given their very low frequencies (e.g., proportion of unsuccessful uptake). However, the RM ANOVA analysis was used for all variables, since there is no non-parametric alternative to the whole design (Field, 2005; Larson-Hall, 2010). Regarding Sphericity, the repeated measures variable in the current study had only two levels (the problem-solving task and the debate task), so the assumptions of Sphericity using *Mauchly's* test were not a concern (Field, 2005; Larson-Hall, 2010). Where there was a statistically significant main effect of proficiency, Tukey HSD post hoc analysis was further reported to see the difference between dyad groups. The Statistical Package for the Social Sciences (SPSS 18.0) was used as the statistical tool to analyse the data. The alpha level was conventionally set at .05. Given that some scholars even argue the alpha should be elected at .10 for social sciences (Kline, 2004; Larson-Hall, 2010; Murphy & Myors, 2004), for those effects that approached significance, .10, and with large effect sizes, further analysis was done to examine where the difference might lie because “the probabilities of Type I and Type II errors are a trade-off, and if the

probability of a Type I error is lowered then the probability of a Type II error is increased” (Larson-Hall, 2010, p.102). As for the RM ANOVA results, percentage variance effect sizes indexed by partial eta squared (η_p^2) values and observed power were also reported (Larson-Hall, 2010). η_p^2 values of .01, .06, and .14 are considered small, medium and large respectively (Sink & Stroh, 2006). The effect size, calculated by correlation coefficient (r) was also reported for paired samples or non-parametric Wilcoxon signed rank tests where they were further used. r values of .20, .30 and .50 are considered small, medium and large respectively (Cohen, 1988). Correlation tests both parametric and non-parametric were also used as appropriate to explore the links between LRE resolutions in task rehearsal and uptake in task performance.

7.7.8 Qualitative analysis

In addition to the main quantitative data, qualitative analysis of episodes of rehearsal and performance and students’ interviews was also used to add more explanatory power to the interpretation of the findings (Dörnyei, 2007; Ellis & Barkhuizen, 2005; Hashemi, 2012).

7.8 Validity and reliability

7.8.1 Validity

Validity addresses “the accuracy of the findings” (Creswell, 2009, p.215) or whether the study “measures what it is supposed to measure” (Dörnyei, 2007, p.50). In the current study, great effort was made to maximise the validity of the research. First, the study adopted a mixed design, with task type (the problem-solving task and the debate) as a within-subject variable and proficiency as a between-subject variable. With this design, reasonable claims could be made about the main effects of tasks and proficiency and the interaction between them. Second, the study held the topic of the two tasks and the time for task rehearsal constant, thus reducing the confounding effects, if any of the differing topics or amounts of rehearsal time on learning behaviours. Next, the research adopted careful procedures for selecting students: (1) it selected students of the same Grade 11 level; (2) it used complementary objective and subjective methods (English scores and teachers’ ratings); (3) it employed randomisation of selected students into proficiency groups (HH, HL, LL) in each class;

(4) it used an equal sample size for each dyad group ($n= 8$); and (5) it had almost similar distributions of gender for each dyad group.

In addition, the piloting, familiarisation sessions, and the counterbalancing of the order of tasks and task performance added more validity to the research. The study also took into consideration of the 'Hawthorne' effect (Dörnyei, 2007; McDonough & McDonough, 1997) or the 'observer's paradox' (Labov, 1972, cited in Nunan, 1992) that the mere presence of the researcher in the classroom might cause the teachers and students to behave differently from what they normally would. In this study, students had gone through several familiarisation sessions prior to the main data collection. Therefore, by the time the main data were collected, the students were used to my presence in the classroom. Furthermore, when I was in the classroom, I sat unobtrusively and quietly in a side corner at the back of the class. Above all, the students understood that I would not evaluate them nor their teachers. It was observed that they behaved naturally with chatting, laughter, and considerable use of L1 in their task rehearsal and relaxed acting with laughter and teasing in performance. The teachers were relaxed and natural in their behaviour, as the focus was on students rehearsing and performing new communicative tasks. Furthermore, the teachers were my colleagues and we were of the same status, and they knew that I was not making observations for assessment. Next, the validity of the research was assured through careful briefings with the teachers before the tasks were carried out so as to ensure consistency in the implementation procedures in all the classes. Finally, the research was conducted in intact classes in their normally scheduled classroom hours, thus maintaining the ecological validity (Loewen & Philp, 2012). Essentially, the study was contextualised in an on-going new task-based English curriculum for high school students in Vietnam, hence addressing the 'pressing' call for contextualised task research with a classroom focus (Bygate, 2001; Bygate et al., 2009; Samuda & Bygate, 2008).

7.8.2 Reliability

The study was believed to satisfy the reliability standards in different ways. First, the study was explicit in every step of data collection. In terms of analysis, the study used

explicit layered coding, that is, one focus each time. For example, LREs were identified and tested for inter-reliability first. Then these LREs were categorised whether they were lexical or grammatical LREs and how they were resolved: correct, incorrect, and unresolved. Subsequently uptake was traced back systematically to LRE resolutions in task rehearsal. Furthermore, the definitions of LREs and their resolutions used in the current study were also attested in previous studies and were further described in detailed coding protocols as they were applied to the current data. This helped enhance the transparency and consistency of the coding. Next, the study used narrow uptake categories, which were 'as narrow as possible' (Révész, 2012), and linked back to the LRE resolutions (e.g., incorrect uptake = incorrect use of incorrectly solved items; unsuccessful uptake = incorrect use of correctly solved items, etc.). Furthermore, the process-product approach (i.e., LREs as learning opportunities created in task rehearsal and the uptake of them in performance), used in coding uptake increased the accuracy in evidencing uptake. Finally, the study formally tested for inter-coder reliability. The data selected for independent coding were equally representative of the two tasks and three dyad groups with the proficiency of the participants 'blinded'. The reliability scores, both simple percentage agreement and Cohen's kappa values showed satisfactory levels of reliability.

In brief, the study was believed to achieve high standards of validity and reliability through the careful design of the study, and carefully planned and executed procedures for the data collection, analysis and reliability.

7.9 Summary

The study adopted a mixed design to examine the impact of proficiency and task type on the learner generation of LREs and how they transferred their language solutions arrived at in rehearsal to performance. Data collection took place in normally scheduled classrooms with the learners' usual teachers. The total data comprised 48 rehearsals and of course 48 corresponding performances, 16 (8x 2 tasks) for each proficiency group (HH, HL, LL). First, rehearsal talk was transcribed and coded for LREs. The LREs identified were then coded into lexical and grammatical LREs, and how they were resolved: correctly solved, incorrectly solved and unsolved. The language items

focussed on in the LREs were used as 'potential tracers' (Storch, 2002b) to trace their take-up in task performance. Uptake was categorised into different types: successful uptake, unsuccessful uptake, incorrect uptake and no uptake. The uptake types were then linked back to how LREs were resolved in task rehearsal. Qualitative analysis of relevant episodes and student interview data were additionally used to complement the quantitative findings. The reliability and validity of the research were taken into account through careful study design including selection of dyads, tasks, task procedures, familiarisation sessions, and piloting. Inter-reliability was also checked and yielded satisfactory levels. The following chapter, Chapter 8 will first present and discuss the findings related to the effects of task type and proficiency on the frequency of LREs and LRE resolutions in rehearsal before the results on the uptake in performance are reported in Chapter 9.

Chapter 8 THE PHASE 2 STUDY: LRES IN TASK REHEARSAL

8.1 Introduction

We have seen from Phase 1 that the teachers in the current EFL context employed a rehearsal-(public) performance approach to implement tasks. They believed that performance had catalysing power in pushing learners to use the target language and engaging them in task work. The students also valued the notion of performance and an opportunity to rehearse for that performance (see Chapters 5, 6). Phase 2 followed up on Phase 1 to specifically address the impact of task design and proficiency on learning opportunities, that is, LREs that arose in task rehearsal and uptake of them in task performance. This chapter focuses on how rehearsal provided opportunities for dyads of differing proficiency levels to problem-solve their meaning-making (i.e., discuss LREs) in two tasks: one problem-solving task and one debate task in preparation for their public co-performance. The chapter reports on and discusses the findings related to four research questions (RQ):

- RQ1. Did tasks and proficiency impact on the occurrences of LREs in task rehearsal?
- RQ2. How did Vietnamese EFL learners resolve their LREs in rehearsal?
- RQ3. Did tasks and proficiency impact on how learners resolved correctly, incorrectly or left unsolved LREs in rehearsal?
- RQ4. Did the linguistic focus of LREs affect how learners resolved them in rehearsal?

The chapter has three main sections. Section 8.2 presents the findings for each research question. Section 8.3 discusses the results and Section 8.4 provides a summary of the chapter.

8.2 Findings

8.2.1 Did tasks and proficiency impact on the occurrences of LREs in task rehearsal?

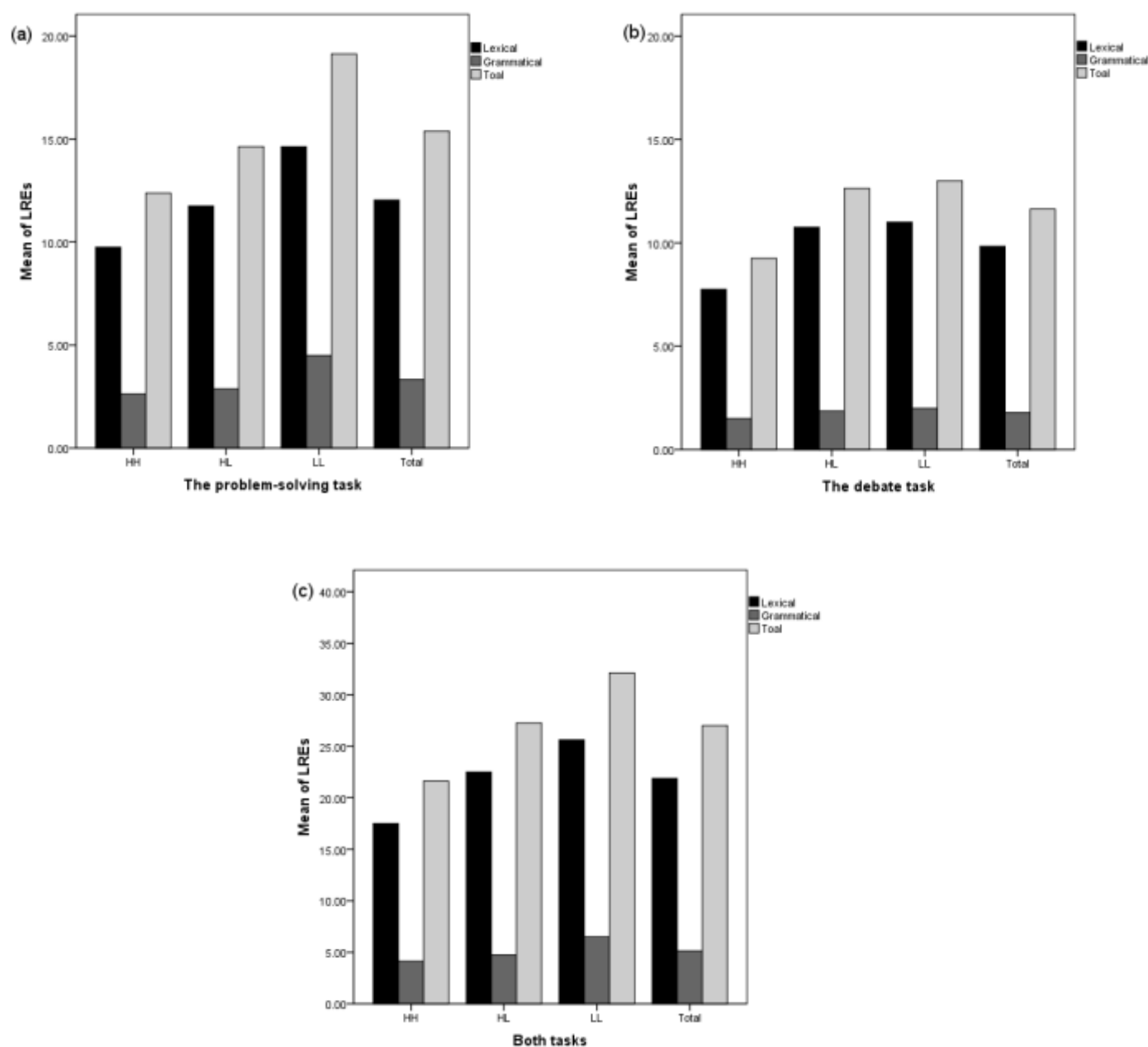
Table 8.1 shows LREs across tasks and proficiency groups which are visually displayed in Figure 8.1.

TABLE 8.1: LREs by task and proficiency

		HH(n= 8 dyads)			HL(n= 8 dyads)			LL(n= 8 dyads)			Total		
		<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>
T 1	Lexical	78	9.75	4.02	94	11.75	2.66	117	14.63	5.34	289	12.04	4.47
	Grammatical	21	2.63	1.06	23	2.88	2.36	36	4.50	2.50	80	3.33	2.16
	<i>Total</i>	99	12.38	4.34	117	14.63	4.03	153	19.13	6.96	369	15.38	5.80
T2	Lexical	62	7.75	3.92	86	10.75	3.62	88	11.00	3.20	236	9.83	3.75
	Grammatical	12	1.50	1.31	15	1.87	1.80	16	2.00	1.85	43	1.79	1.61
	<i>Total</i>	74	9.25	4.13	101	12.63	4.17	104	13.00	4.34	279	11.63	4.38
Total	Lexical	140	17.50	7.58	180	22.50	4.41	205	25.63	6.41	525	21.88	6.90
	Grammatical	33	4.13	1.73	38	4.75	3.01	52	6.50	3.02	123	5.13	2.74
	<i>Total</i>	173	21.63	8.16	218	27.25	5.90	257	32.13	9.13	648	27.00	8.68

Note. T1= the problem-solving task; T2= the debate task

FIGURE 8.1: LREs cross tasks and proficiency groups



As is seen from Table 8.1 and Figure 8.1, there is a tendency for all proficiency groups to generate more LREs when rehearsing the problem-solving task than the debate task. In addition, as proficiency in the dyads decreases, the occurrences of LREs tend to increase. This trend is apparently true not only for the total data, but also for each individual task. With regards to the linguistic focus of LREs (lexical/grammatical), a lexical focus is predominant for both tasks, all proficiency groups and in all task rehearsals.

In order to test the significance of these trends, an RM ANOVA was conducted on the data with two within-subject variables, namely, task (with two levels: the problem-solving task and the debate task) and linguistic focus (lexical and grammatical), and one between-subject variable, proficiency (with three levels: HH, HL, LL). Table 8.2 summarises the results of the RM ANOVA for LREs across tasks, linguistic foci and proficiency.

TABLE 8.2: Summary of the RM ANOVA results on LREs across task type, linguistic focus and proficiency

Source		<i>F</i>	<i>p</i>	η_p^2	Observed <i>power</i>
<i>Within-subjects</i>	Task type	11.363	.003	.35	.90
	Linguistic focus	215.472	.000	.91	1.00
	Task * proficiency	1.225	.314	.10	.24
	Linguistic focus * proficiency	2.308	.124	.18	.42
	Task* Linguistic focus	.399	.535	.02	.09
	Task * Linguistic focus *	.104	.901	.01	.06
	Proficiency				
<i>Between-subjects</i>	Proficiency	3.589	.046	.26	.60

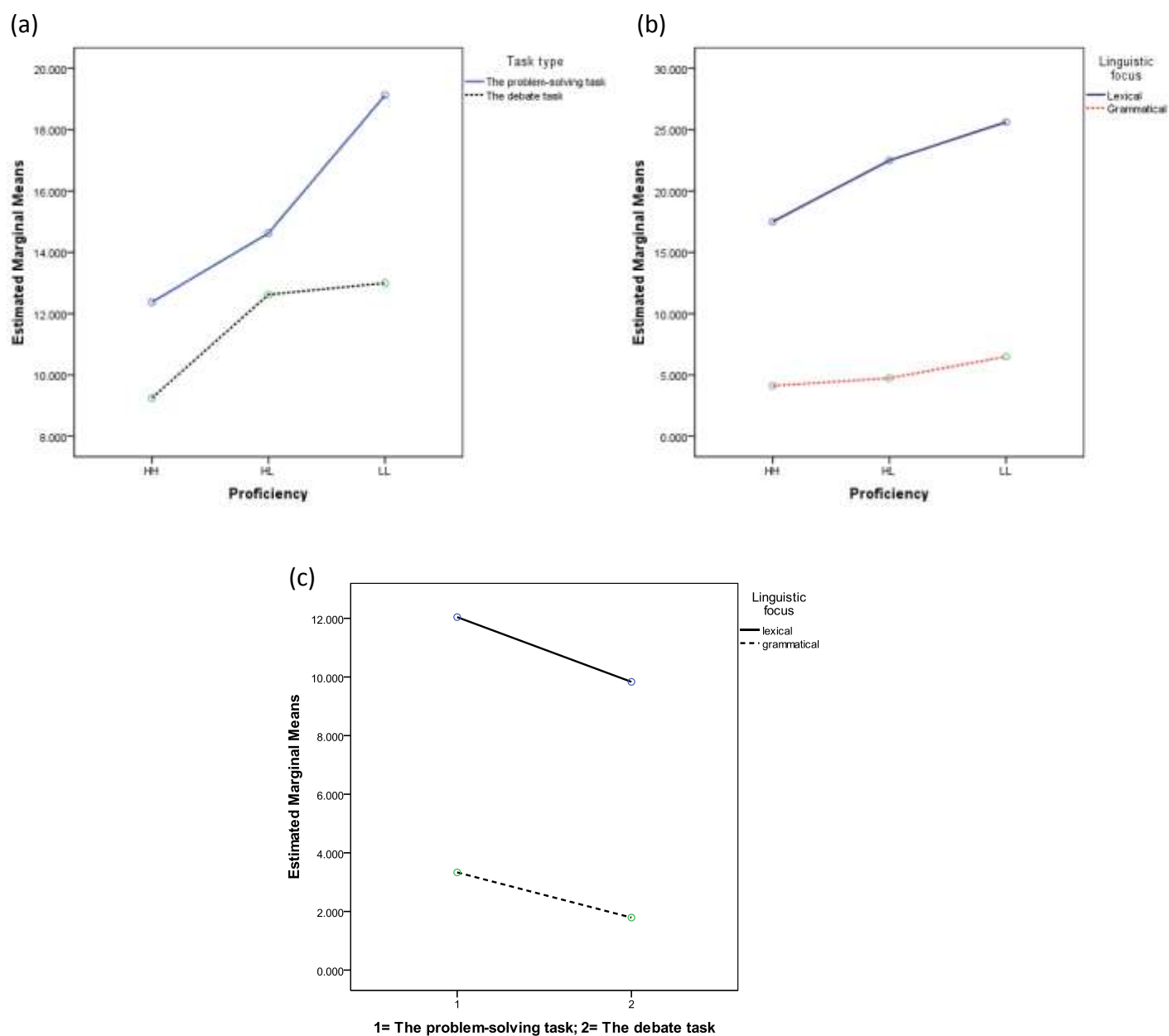
As Table 8.2 shows, there was a statistically significant main effect of task, $F(1,21)=11.363$, $p=.003$, $\eta_p^2=.35$, suggesting that learners generated significantly more LREs when rehearsing the problem-solving task ($M=15.38$, $SD=5.80$) than the debate task ($M=11.63$, $SD=4.38$). There was also a significant main effect of linguistic focus, $F(1, 21)=215.472$, $p<.001$, $\eta_p^2=.91$, indicating that learners, regardless of proficiency and task type, paid central attention to vocabulary when discussing their language problems. Dyads discussed, on average, 21.88 lexical LREs ($SD=6.90$) and 5.13 grammatical LREs ($SD=2.74$).

The RM ANOVA results also showed a significant main effect of proficiency, $F(2, 21)=3.589$, $p=.046$, $\eta_p^2=.26$. The mean number of LREs for the dyad groups was: for the HH, $M=21.63$ ($SD=8.16$); for the HL, $M=27.25$ ($SD=5.90$) and for the LL, $M=32.13$ ($SD=9.13$). Tukey HSD post-hoc analysis revealed that the LL dyads generated

significantly more LREs than the HH ($p=.036$), but the difference between the HH and HL and between the HL and LL was not statistically significant ($p=.342$ and $p=.442$ respectively). In other words, the effects of the lowest proficiency prevailed.

None of the interaction effects were statistically significant: between task and proficiency, $F(2, 21)=1.225$, $p=.314$, between linguistic focus and task, $F(2, 21)=.399$, $p=.535$, between linguistic focus and proficiency, $F(2,21)=2.308$, $p=.124$, and between task, proficiency and linguistic focus, $F(2, 21)=.104$, $p=.901$. All these suggest that i) dyads consistently produced more LREs when rehearsing for the performance of the problem-solving task than for the debate task ii) preference for lexical LREs rather than grammatical LREs was consistent throughout both tasks and all proficiency groups. All these effects can be visually seen in Figure 8.2.

FIGURE 8.2: LREs by task, proficiency and linguistic focus



In summary, learners, irrespective of proficiency groups, discussed significantly more LREs in the problem-solving task than the debate task. Learners paid much more attention to vocabulary than grammar regardless of tasks and proficiency. With decreased overall dyad proficiency, they tended to discuss more LREs. However, the difference was significant only between the LL and HH dyads.

8.2.2 How did Vietnamese EFL high school learners resolve language problems in task rehearsal?

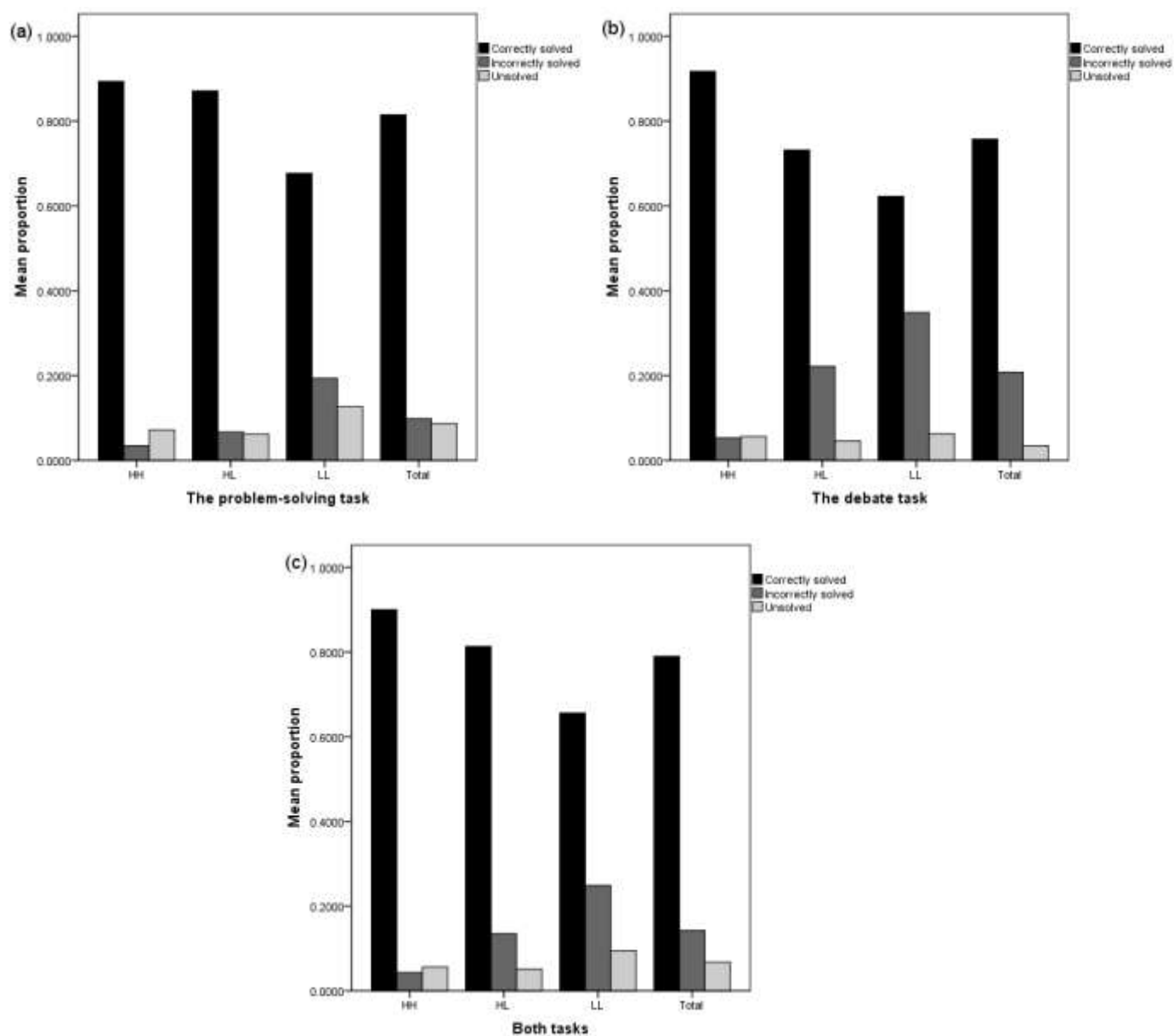
In posing this research question, I was interested in how the learners resolved language problems they were faced with in task rehearsal. LREs were resolved in three ways: correctly solved, incorrectly solved and unsolved. LRE resolutions were calculated as proportion scores for each type of resolution out of the total LREs. Table 8.3 summarises LRE resolutions by task and proficiency which are visually displayed in Figure 8.3.

TABLE 8.3: LRE resolutions by task and proficiency

		HH (n= 8 dyads)			HL (n= 8 dyads)			LL(n= 8 dyads)			Total		
		<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>
Correctly solved	T1	88	.89	.09	102	.87	.12	100	.68	.12	290	.81	.14
	T2	67	.92	.12	74	.73	.11	64	.62	.15	205	.76	.17
	<i>Total</i>	155	.90	.08	176	.81	.08	164	.66	.10	495	.79	.13
Incorrectly solved	T1	4	.03	.05	8	.07	.09	31	.19	.09	43	.10	.10
	T2	4	.05	.09	22	.22	.10	36	.35	.16	62	.21	.17
	<i>Total</i>	8	.04	.06	30	.13	.07	67	.25	.07	105	.14	.11
Unsolved	T1	7	.07	.07	7	.06	.08	22	.13	.10	36	.09	.09
	T2	3	.03	.05	5	.05	.06	4	.03	.06	12	.03	.06
	<i>Total</i>	10	.06	.04	12	.05	.04	26	.09	.06	48	.07	.05
Frequency of LREs	T1	99	12.38	4.34	117	14.63	4.03	153	19.13	6.96	369	15.38	5.80
	T2	74	9.25	4.13	101	12.63	4.17	104	13.00	4.34	279	11.63	4.38
	<i>Total</i>	173	21.63	8.16	218	27.25	5.90	257	32.13	9.13	648	27.00	8.68

Note. T1= The problem-solving task; T2= The debate task

FIGURE 8.3: LRE resolutions across tasks and proficiency groups



As is seen from Table 8.3 and Figure 8.3, the Vietnamese EFL learners working collaboratively in rehearsal were largely successful at resolving their LREs. All proficiency groups correctly resolved the majority of their LREs than incorrectly resolved them or abandoned them unsolved.

8.2.3 Did task type and proficiency affect how LREs were resolved in task rehearsal?

This research question addresses whether task type and proficiency played a role in how the learners resolved their LREs. To test the effects of task type and proficiency on the resolutions of LREs, separate RM ANOVAs with one within-subject variable- task type (with two levels: the problem-solving and the debate task) and one between-subject variable: proficiency (with three levels: HH, HL, LL) were carried out with each type of LRE resolution as the dependent variable. Table 8.4 summarises these RM ANOVA results.

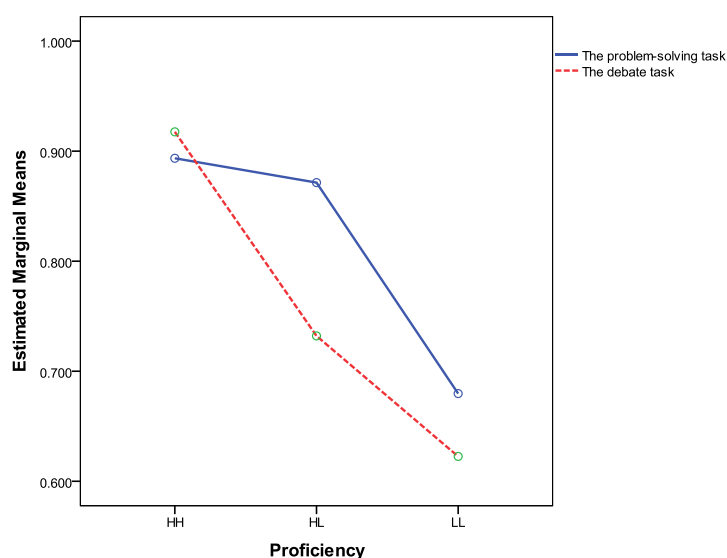
TABLE 8.4: Summary of the RM ANOVA results for LRE resolutions

Source		Measure	<i>F</i>	<i>p</i>	η_p^2	Observed power
<i>Within-subjects</i>	Task	Correctly solved LREs	3.717	.068	.15	.45
		Incorrectly solved LREs	18.210	.000	.46	.98
		Unsolved LREs	4.634	.043	.18	.54
	Task type* Proficiency	Correctly solved LREs	2.501	.106	.19	.45
		Incorrectly solved LREs	3.161	.063	.23	.54
		Unsolved LREs	1.010	.381	.09	.20
<i>Between-subjects</i>	Proficiency	Correctly solved LREs	15.677	.000	.60	1.00
		Incorrectly solved LREs	16.679	.000	.61	1.00
		Unsolved LREs	.878	.430	.08	.18

Regarding the measure of correctly solved LREs, the results showed no statistically significant main effect of task type, $F(1, 21)=3.717$, $p=.068$, $\eta_p^2=.15$, even though learners correctly solved a higher proportion of LREs in the problem-solving task ($M=.81$, $SD=.14$) than the debate task ($M=.76$, $SD=.17$). However, there was a significant main effect of proficiency, $F(2, 21)=15.677$, $p<.001$. This was a large effect size ($\eta_p^2=.60$). The mean proportions of correctly solved LREs for the HH, the HL and the LL are .90 ($SD=.08$), .81 ($SD=.08$), and .66 ($SD=.10$) respectively. Tukey HSD post-hoc analysis showed that both the HH and the HL correctly solved a significantly higher proportion of LREs than the LL dyads ($p<.001$ and $p=.009$ respectively). There was no

statistically significant difference between the HH and the HL groups ($p=.082$). However, the results showed no significant interaction effect between task and proficiency, $F(1,21)=2.501$, $p=.106$, indicating that the impact of proficiency was consistent for the two tasks (see Figure 8.4).

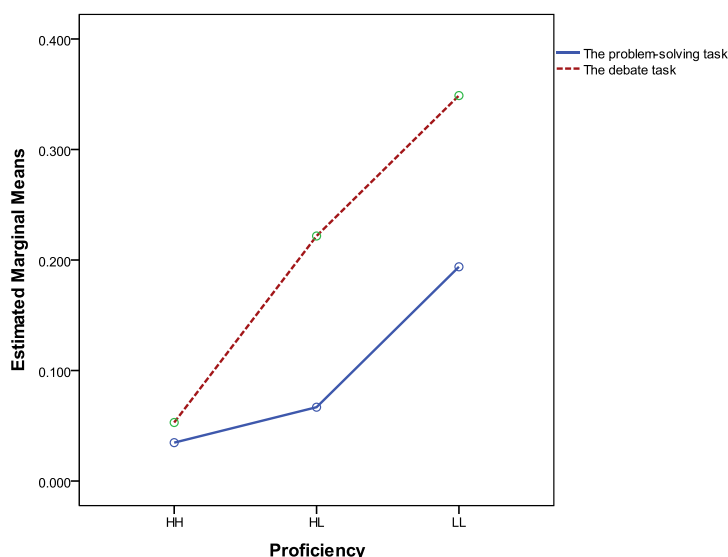
FIGURE 8.4: Correctly solved LREs by task and proficiency



For the incorrectly solved LREs, the results showed a significant main effect of task, $F(1, 21)=18.210$, $p<.001$. This was a large effect size ($\eta_p^2=.46$), indicating that learners incorrectly solved a significantly higher proportion of LREs in the debate task ($M=.21$; $SD=.17$) than in the problem-solving task ($M=.10$, $SD=.10$). There was also a significant main effect of proficiency, $F(2, 21)=16.679$, $p<.001$, and this was a very large effect size ($\eta_p^2=.61$). The mean proportions of incorrectly resolved LREs for the dyad groups were $M=.04$ ($SD=.06$) for the HH, $M=.13$ ($SD=.07$) for the HL and $M=.25$ ($SD=.07$) for the LL. Tukey HSD post-hoc analysis showed all dyad groups differed statistically from one another. Specifically, there was a statistically significant difference between the HH and the LL groups ($p<.001$), between the HL and the LL ($p=.011$), and between the HH and the HL ($p=.048$). Overall, the LL group had the greatest proportion of incorrect resolutions, followed by the HL and then the HH. There was a non-significant interaction effect between task and proficiency, though it was approaching significance, $F(2, 21)=3.161$, $p=.063$, $\eta_p^2=.23$, indicating all dyad groups, especially the

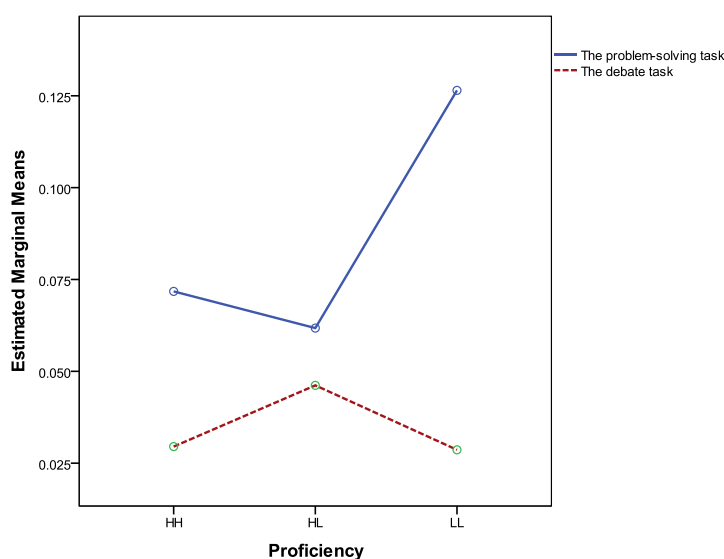
HL and LL, had a higher proportion of their LREs incorrectly solved in the debate task than the problem-solving task (see Figure 8.5).

FIGURE 8.5: Incorrectly solved LREs by task and proficiency



As far as unsolved LREs are concerned, as Table 8.4 shows, there was a significant main effect of task, $F(1, 21) = 4.634, p = .043$. A greater proportion of LREs was unsolved in the problem-solving task ($M = .09, SD = .09$) than the debate task ($M = .03, SD = .06$). This task effect did not differ by proficiency, as shown by the non-significant interaction effect between task and proficiency, $F(2, 21) = 1.010, p = .381$ (see Figure 8.6). The main effect of proficiency, was not statistically significant, either, $F(2, 21) = .878, p = .430$, indicating that leaving LREs unsolved was not influenced by whether the dyads were HH, HL, or LL.

FIGURE 8.6: Unsolved LREs by task and proficiency



In short, success at resolving language problems in task rehearsal was influenced by learner proficiency and tasks. Lower proficiency dyads were less successful at correctly resolving their LREs than higher proficiency dyads. However, once they left their LREs unsolved, this was not affected by proficiency. Tasks did not have a significant effect on students correctly resolving language problems. Yet they problem-solved incorrectly more in the debate task than in the problem-solving task, although they left a higher proportion of their LREs unsolved in the latter task.

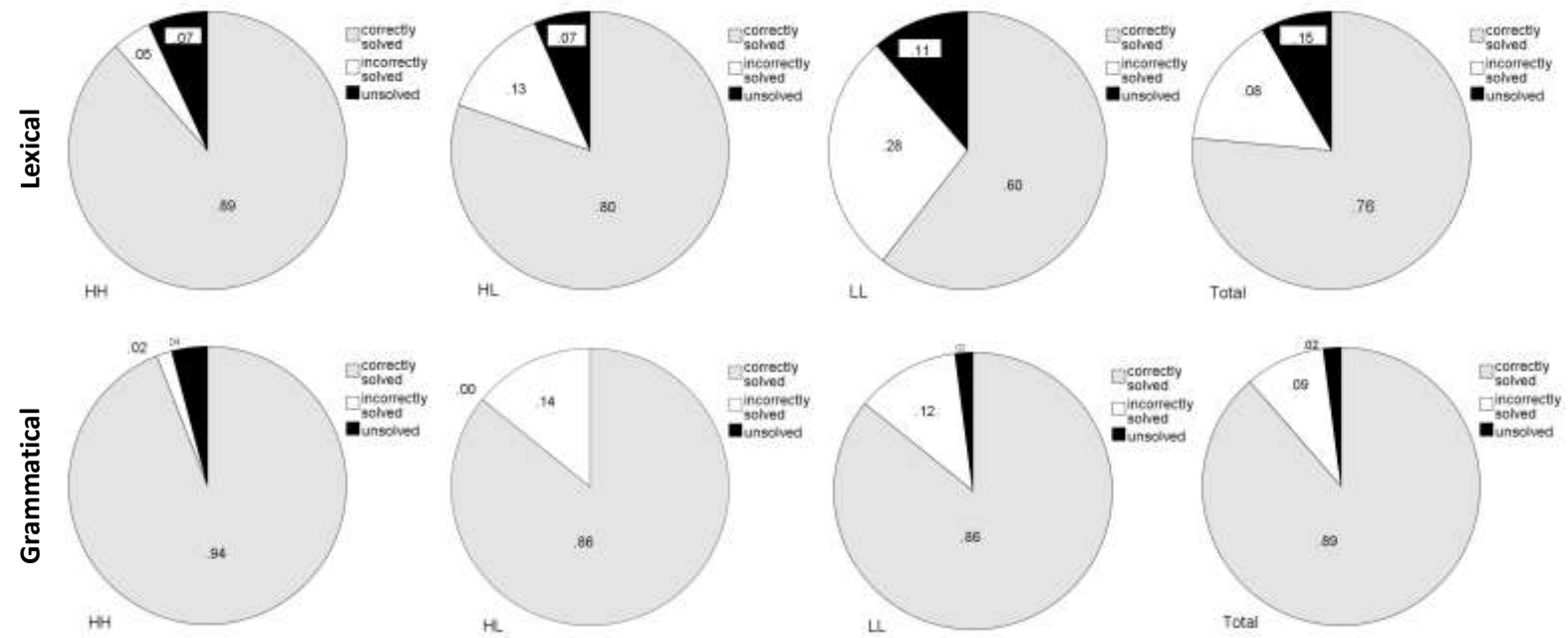
8.2.4 Did the linguistic focus of LREs affect how learners resolved them in task rehearsal?

This research question is concerned with whether or not learners resolved their lexical and grammatical LREs differently. In order to answer this question, the data of the two tasks were merged. This was done because the much lower frequency of grammatical LREs relative to lexical LREs meant that splitting them by task would reduce the frequency to a level that could not be reliably analysed using inferential statistics. The proportion of lexical or grammatical LREs being correctly solved, incorrectly solved or unsolved was calculated out of the total lexical or grammatical LREs. Table 8.5 provides data on LRE resolutions by linguistic focus and proficiency which are further displayed in Figure 8.7.

TABLE 8.5: LRE resolution by linguistic focus

	Correctly solved		Incorrectly solved		Unsolved		Frequency of LREs	
	<i>Lexical</i>	<i>Grammatical</i>	<i>Lexical</i>	<i>Grammatical</i>	<i>Lexical</i>	<i>Grammatical</i>	<i>Lexical</i>	<i>Grammatical</i>
HH (n= 8 dyads)								
<i>Sum</i>	124	31	7	1	9	1	140	33
<i>M</i>	.89	.94	.05	.02	.07	.04		
<i>SD</i>	.09	.12	.07	.05	.06	.12		
HL (n= 8 dyads)								
<i>Sum</i>	144	32	24	6	12	0	180	38
<i>M</i>	.80	.86	.13	.14	.07	.00		
<i>SD</i>	.08	.18	.08	.18	.06	.00		
LL(n= 8 dyads)								
<i>Sum</i>	121	43	60	7	24	2	205	52
<i>M</i>	.60	.86	.28	.12	.11	.02		
<i>SD</i>	.11	.15	.09	.11	.07	.06		
Total								
<i>Sum</i>	389	106	91	14	45	3	525	123
<i>M</i>	.76	.89	.15	.09	.08	.02		
<i>SD</i>	.16	.15	.13	.13	.04	.07		

FIGURE 8.7: LRE resolution by linguistic focus



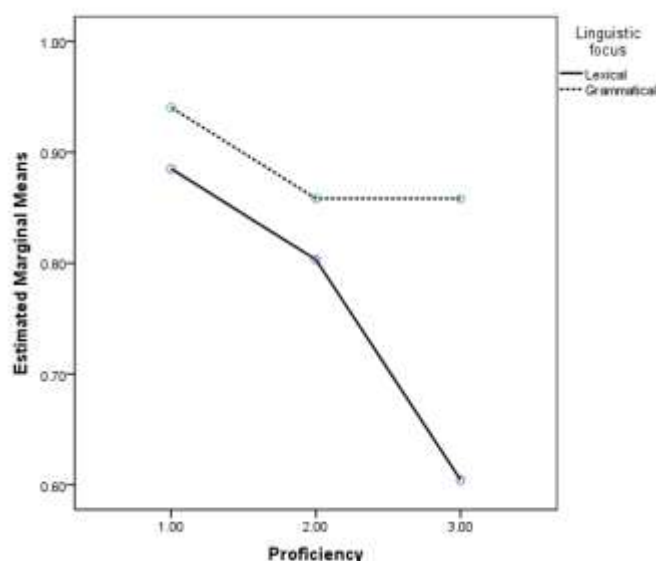
As Table 8.6 and Figure 8.7 show, learners seemed to resolve lexical and grammatical problems differently. In order to test the significance of the main trends in the data, separate RM ANOVAs with one within-subject variable (linguistic focus: lexical and grammatical) and one between-subject variable (proficiency: HH, HL, LL) were performed with each type of LRE resolution as the dependent variable. The RM ANOVA results were provided in Table 8.6.

TABLE 8.6: Summary of the RM ANOVA results for LRE resolutions by linguistic focus and proficiency

Source		Measure	<i>F</i>	<i>p</i>	η_p^2	Observed power
Within-subjects	Linguistic focus	Correctly solved LREs	16.421	.001	.44	.97
		Incorrectly solved LREs	4.324	.050	.17	.51
		Unsolved LREs	9.305	.006	.31	.83
	Linguistic focus*	Correctly solved LREs	4.881	.018	.32	.74
		Incorrectly solved LREs	3.327	.056	.24	.57
	Proficiency	Unsolved LREs	.857	.439	.08	.18
Between-subjects	Proficiency	Correctly solved LREs	6.073	.008	.37	.84
		Incorrectly solved LREs	9.240	.001	.47	.96
		Unsolved LREs	1.021	.377	.09	.20

Regarding the question of whether lexical or grammatical LREs were more likely to be correctly solved, the results showed a significant interaction between linguistic focus and proficiency, $F(2,21)=4.881$, $p=.018$, and the effect size was large ($\eta_p^2=.32$), suggesting that the effect of linguistic focus on the likelihood of successful resolution of LREs was mediated by proficiency (see Figure 8.8).

FIGURE 8.8: Correctly solved LREs by linguistic focus

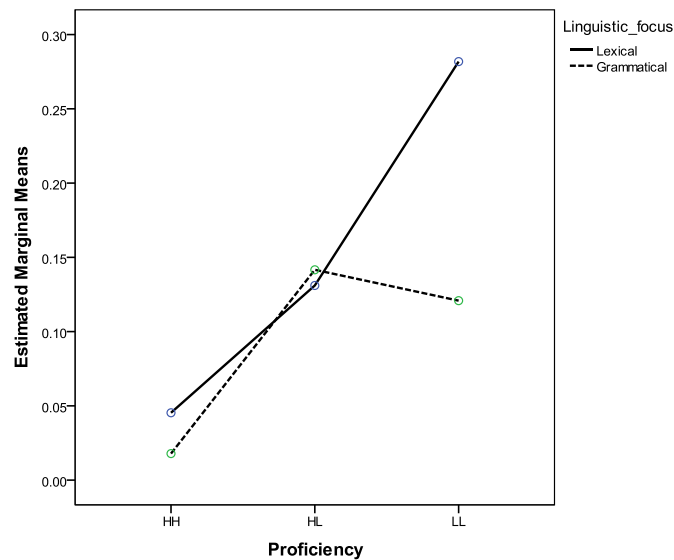


Individual paired samples *t*-tests or non-parametric Wilcoxon Signed Ranks tests further showed that the difference between the proportions of lexical and grammatical LREs being correctly solved was not significant for the HH group ($Z=-1.153$, $p=.249$, $r=.06$) and for the HL, $t(7)=-1.027$, $p=.339$, $r=.36$. However, the LL dyads correctly resolved a significantly greater proportion of grammatical ($M=.86$, $SD=.15$) than lexical ($M=.60$, $SD=.11$) LREs, $t(7)=-5.276$, $p=.001$, $r=.89$. These results suggest that while the HH and HL dyads were able to correctly solve lexical and grammatical LREs equally well, the LL did significantly better when the focus was on grammar.

With regard to whether lexical or grammatical LREs were more likely to be incorrectly solved, the results showed the interaction between linguistic focus and proficiency was very close to significance, $F(2,21)=3.327$, $p=.056$. However, since the effect size was large ($\eta_p^2=.24$), further individual paired samples *t*-tests or non-parametric Wilcoxon Signed Ranks tests as appropriate were carried out. The results showed that for the HL dyads, the difference between the proportions of lexical ($M=.13$, $SD=.08$) and grammatical ($M=.14$, $SD=.18$) LREs being incorrectly resolved was not significant, $t(7)=-.162$, $p=.876$, $r=.06$. The LL dyads incorrectly solved a significantly higher proportion of lexical ($M=.28$, $SD=.09$) than grammatical ($M=.12$, $SD=.11$) LREs, $t(7)=3.330$, $p=.013$, $r=.78$. Although the HH dyads incorrectly resolved a higher proportion of lexical

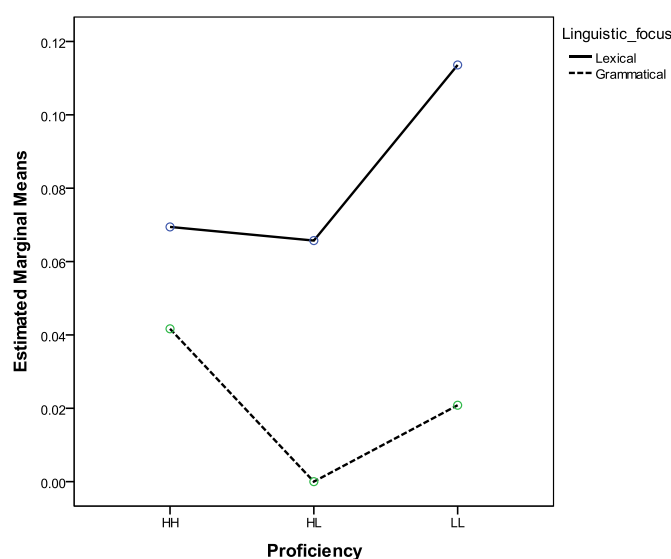
($M=.05$, $SD=.07$) than grammatical ($M=.02$, $SD=.05$) LREs, the difference was not significant, $Z=-.730$, $p=.465$, $r=-.18$. Overall, when two lower proficiency learners worked together, they were more likely to incorrectly resolve lexical than grammatical LREs (see Figure 8.9).

FIGURE 8.9: Incorrectly solved LREs by linguistic focus



For the unsolved LREs, the results showed a significant main effect of linguistic focus, $F(1,21)=9.305$, $p=.006$. The effect size was large ($\eta_p^2=.31$), with a significantly higher proportion of unsolved lexical ($M=.08$) than grammatical ($M=.02$) LREs. The non-significant interaction between linguistic focus and proficiency, $F(2, 21)=.857$, $p=.439$, further showed that lexical LREs were more likely to be left unsolved than grammatical LREs for all proficiency groups (see Figure 8.10).

FIGURE 8.10: Unsolved LREs by linguistic focus



In short, the HH and HL dyads correctly resolved both lexical and grammatical LREs equally well, while the LL dyads were more successful when their LREs dealt with grammar. Also only the LL dyads incorrectly solved a significantly higher proportion of their lexical than grammatical problems, though all proficiency groups consistently left a higher proportion of lexical than grammatical LREs unsolved.

8.3 Discussion

As shown in Phase 1, in these Vietnamese EFL classrooms, the teachers preferred open-ended tasks, particularly divergent over convergent tasks in order to engage students socio-affectively and maximise output. Decision making by the teachers reflected a general commitment to a final *public performance* of the task by groups of students (see Chapter 5). This public performance was preceded by *rehearsal* for the performance, involving students doing the task in pairs or groups to prepare for the performance of it in front of the class. Phase 2 examined how tasks and proficiency impacted on opportunities for learning in a rehearsal-performance condition. The findings have shown that given an opportunity to rehearse for their public performance, students frequently attended to their language use (i.e., discussed LREs) and that tasks and proficiency affected not only the occurrences of LREs but also how they were resolved. The focus of this discussion will be on (1) rehearsal and the effects

of tasks; (2) rehearsal and the effects of proficiency and (3) rehearsal and the linguistic focus of LREs and LRE resolutions.

8.3.1 Rehearsal and task effects on LRE occurrences and resolutions

The findings showed that learners produced significantly more LREs in rehearsing for the problem-solving task (convergent task) than the debate task (divergent task). This seems to be in line with the assumption that convergent tasks are more likely to induce interaction (Pica et al., 1993) and with research which has found more negotiation of meaning in convergent tasks (Duff, 1986; Nakahama et al., 2001). However, two points need clarification here. First, the current rehearsal data showed it was not so much the convergence or divergence of task goals assumed in convergent and divergent tasks that caused a difference in the number of LREs between the two tasks. For example, upon approaching the problem-solving task, dyad members typically selected two charity options quickly, and collaboratively sought reasons for such choices without going through a process of considering, eliminating, and making final decisions (Episode 1). Similarly, in the debate task, they quickly negotiated who would be for or against the statement and then carried on with the task (Episode 2). In this manner, the convergence presumed in the problem-solving task did not differ markedly from the divergence of the debate task. The learners collaboratively worked towards their joint public performance. The problem was thus how to do the performance itself. Therefore, it may be premature to attribute more LREs in the problem-solving task to its goal-convergent nature.

Episode1 (The problem-solving task-LL-11a)

- S1: Tau nghĩ là nên cho học bổng.
(*I think we should provide scholarships.*)
- S2: Học sinh cấp ba nghèo với là trái tim cho em. Tau chọn hai cái nớ, mi chọn chi?
(*Poor high school students and the heart for children. I choose these two, what about you?*)
- S1: Hai cái nớ hi, 1-3.
(*Yeah, these two, 1-3.*) [options 1 and 3]
- S1: 1-3. Chừ đưa ra lý do.
(*One-three. Now let's give reasons.*)

Episode 2 (The debate task- HH-11c)

- S1: Mi đồng ý hay không đồng ý?
(*Are you for or against?*)
- S2: Phản đối (.) phản đối khó hơn (*Against (.) against is harder.*)
- S1: Tau đồng ý. (*I am for.*)
- S2: Rồi. (*Ok.*)
- S1: Tau đồng ý, mi against. Lại tình huống chớ.
(*I am for, you are against. Think of situations first.*)
[situations to open and close their public performance]

Second, LREs here did not arise from negotiation of meaning to achieve greater comprehensibility (see Foster & Ohta, 2005, for a detailed discussion of negotiation of meaning). Although focus on form may be “triggered by perceived problems with *comprehension* or *production*” (Long & Robinson, 1998, p.23, italics added), a majority of the LREs students discussed involved production rather than comprehension problems, as exemplified in Episodes 3-5.

Episode 3 (The problem-solving task-HH-11d)

- S1: (...) my family er have to (.) di dời là chi? (*how to say 'move'?*)
- S2: di dời (*move*) (.) ah move!
- S1: move to erm (.) flood shelters.

It is clear that the learners began with a communicative intention in Vietnamese via an explicit request (how to say ‘move’ (in English)) which entailed searching for the relevant English word to encode. Even in co-construction cases, intended meanings were explicit in L1 (Episode 4).

Episode 4 (The debate task-HL-11d)

- S2: what about (.) how about handicapped people (.)
and they don’t (.) they don’t sức lao động ((have) the ability to work) (.) they can’t
được (just use they can’t)
- S1: it’s impossible
- S2: it’s impossible for them to (.) tự nuôi sống bản thân a (.) nuôi sống ...
(I mean, to support themselves (.) support ...)
- S1: nuôi sống là chi hèo... a! support là cung cấp!
(how to say support ... a! support means ‘provide’!)
- S2: Rồi, rằng nữa? (Ok. What else?)

In Episode 4, S2 originally intends to say ‘they don’t have the ability to work’ but he does not seem able to find the relevant English phrase, so he shifts ‘they can’t (work)’. Then S1 proposes another alternative ‘it’s impossible’, which S2 then follows up on. S2 afterwards wants to encode the explicit L1 meaning ‘nuôi sống bản thân’ in English. S1 then assists by using Vietnamese L1 to mediate his search for the solution. Clearly the two interlocutors collaboratively formulate the message ‘it’s impossible for them to support themselves’ (instead of ‘they don’t have the ability to work or they can’t work’), as originally intended, without having any comprehensibility problems. In other words, learners were more occupied with productive rather than receptive vocabulary. ‘A need to mean’ (Samuda, 2001) was the trigger for pushing learners to retrieve, search for and pool English resources to provide appropriate forms of expression. Even in grammatical LREs, learners were not concerned with comprehensibility, either, but instead chose to make their language production more accurate or more target-like, as shown in Episode 5.

Episode 5 (The problem-solving task- LL-11c)

S1: oh that's really a big (.) big money erm I think

S2: much! much! money is not đếm được (*countable*).

S1: much (.) much money (.) much money. I think if you should spend it for yourself, just for yourself, it's so [selfish

S2: [selfish

In Episode 5, S1 uses '*a big (.) big money*', which S2 perceives as incorrect and then not only overtly offers a correct correction '*much*', but also gives metalanguage explanation for it. Clearly grammatical LREs, like lexical counterparts, do not necessarily involve comprehension problems.

Given the production focus of the LREs in the current data, the differential meaning-making demands that the two tasks made on the learners could explain why the problem-solving task induced more LREs than the debate task. The problem-solving task required dyad members to decide on two charity options among the five given and provide justifications (see Appendix 8), thus having a more limited choice of content and then a more limited meaning space. In order to make a certain choice, learners needed to provide rationales within the parameter of the content of that choice. For example, upon justifying why they decided to use the budget on building rehabilitation centres for drug addicts, the rehearsal talk showed that learners talked about 'social burdens', 'social integration', 'criminals', and so forth. Similarly, in order to express why they chose to spend the money on building flood shelters for people in central Vietnam, learners tended to talk about 'destroy', 'wash away', 'move', and so forth. These contents were demanding to encode linguistically, yet learners had less freedom to avoid them. The need to encode these contents unavoidably necessitated discussion of relevant grammar-lexical items (Bygate & Samuda, 2009; Robinson, 2001, 2005, 2011a, 2011b; Skehan, 2009), thus pushing learners to generate more LREs. So it was the actual content constraint that pushed output in the problem-solving task.

In contrast, the debate task was more open³³ in terms of the content to be conveyed (i.e., the content/arguments came from the learners). Accordingly, the learners could easily drop complex ideas and associated challenging linguistic means, thus generating fewer LREs. This finding is consistent with the cognition hypothesis (Robinson, 2001, 2005, 2007, 2011a, 2011b) which predicts more focus on form in complex tasks. That is, when the conceptual demand of a task is higher, it is more likely to 'direct' learners to use L2 words or phrases of relevance necessitated by such a conceptual demand. This is because learners have to map conceptualisation or meaning to L2 form (Levelt, 1989; Robinson, 2011a, 2011b; Slobin, 2003). In other words, the 'communicative pressure' of a task lies in its *field* or task content that triggers the use of relevant vocabulary and grammar (Bygate & Samuda, 2009). Note that Robinson operationalises simple or complex tasks by means of presence or absence of a certain task feature such as +/-here and now or +/-few elements of the same task. In the current study, the problem-solving task was regarded as more demanding than the debate because it entailed complex ideas to encode, but less opportunity for avoidance. The learners also perceived the problem-solving task as more difficult:

I think the task where one is for and one is against is easier because we can think of any ideas to say. The other task is more difficult. I have to talk about these options which are harder to say because we lack vocabulary and we don't have much choice.
(HTKN-11c)

A note in order here is that LREs arose in the rehearsal context where learners had 15 minutes to prepare for their public performance. It is the rehearsal, a form of pre-task planning (Ellis, 2005) that further afforded students opportunities to attend to form to encode their wanted meanings. Robinson (2001, 2005) also contends that the optimal condition for focus on form is a conceptually complex task with opportunities to plan. In this regard, it may not be the constraining content which was more demanding to express in the problem-solving *per se* but a combination of the task and task condition (rehearsal-performance) that pushed learners to produce more LREs in this task (see Skehan, 2009 for a discussion of an interaction between tasks and task conditions).

³³ In task-based interaction literature, open and closed tasks are distinguished by whether or not they require a single correct task solution (Pica et al., 1993), though these two variables are not dichotomous, but one task can be more open or closed than another (Ellis, 2003; Willis, 1996).

This implies task effects should not be seen in isolation, but essentially in the conditions in which tasks are carried out.

To reiterate, the learners in the current data rarely asked each other the meaning of an L2 item (cf. Newton, 2013). That is, they rarely produced *receptive* LREs. This was probably because the teachers made sure they understood the task input before they carried out the tasks. Therefore, it was not so much the greater amount of input in the problem-solving task (see Appendix 8), but rather the input-dependence or content constraint to be expressed in this task that entailed more discussion of relevant L2 forms or pushed more LREs. This is pedagogically interesting as teachers may consider designing task content to push learners' *productive* vocabulary. Future research is warranted as to how degrees of input-dependence of a task may affect the nature of LREs (productive/receptive) and L2 learning in a rehearsal-performance condition.

Although the problem-solving task elicited significantly more LREs than the debate task, it did not have a significantly higher proportion of LREs being resolved correctly. One possible explanation could be that the two tasks shared a common feature, which overrode the complexity difference, and that is, they entailed a public performance after rehearsal. The pressure of an imminent performance in both tasks directed attention to form or accuracy, thus reducing task effects. Willis (1996) argues that public performance "creates a need for accuracy" (p.56) (also see Skehan, 1998; Skehan & Foster, 1997). It could be that, due to the forthcoming joint performance, the learners were catalysed to collaboratively pool each other's resources in order to solve language issues they faced, and therefore, they resolved correctly a large majority of problems, regardless of tasks. However, another possible explanation could be the small data sample. The study involved only one problem-solving task and one debate task, so future research with more tasks of the same type under a rehearsal-performance condition may be more likely to find a statistically significant difference.

The finding that learners were more likely to problem-solve incorrectly in the debate task also deserves attention. This finding seems to support the cognition hypothesis in that the debate task, as the easier of the two, according to the students, would induce less accuracy because learners were less likely to process language deeply.

Furthermore, the debating could push learners towards an oral rebuttal that involved arguments and counter-arguments, and therefore inclining them to attentively respond to forthcoming arguments (fluency), hence reducing opportunities to reflect on language form. In this respect, the finding is consistent with the trade-off hypothesis which argues for a competition between fluency and accuracy (Skehan, 1998, 2009; Skehan & Foster, 2001).

In the case of unsolved LREs, it is not surprising that learners abandoned a higher proportion of their LREs unresolved in the problem-solving task. As argued earlier, in this task there was less flexibility in the message and therefore, learners were challenged with meanings needed to express but for which they did not have adequate L2 resources. In contrast, in the debate task, learners were able to provide their reasons independent of any content given (though of course they had to choose either to agree or disagree, but the justifications were at their disposal), which allowed them to avoid certain ideas that entailed language items they did not know. So in the debate task once learners initiated a language problem, they were more likely to come up with a solution (even though incorrect) rather leave them unsolved. This explains why learners left a lower proportion of their LREs unresolved in the debate task.

8.3.2 Rehearsal and proficiency effects on LRE occurrences and resolutions

The study found that lower proficiency dyads discussed more LREs, but were less successful at resolving them. Regarding the occurrences of LREs, the results broadly corroborate Nassaji's (2010) findings that her ESL beginner class initiated more focus on form than intermediate and advanced classes, though the context here was teacher-learner interaction and proficiency levels were class levels. Dobao (2012) found similar results in his laboratory-based study, with intermediate learners generating more lexical LREs than advanced learners doing the same 'find the difference' task. However, the result strikingly contrasts with other studies (Kim & McDonough, 2008; Leese, 2004; Watanabe & Swain, 2007; Williams, 1999) which found more LREs with higher proficiency learners. One possible explanation for these contradictory findings is the different nature of the tasks used. For example, Williams's study covered a wide range of activities from meaning-focussed to form-focussed

activities and learners' differing orientations to these activities could have been an influence. In all the other studies, the dictogloss activities were used. Since the dictogloss tasks involved a given content to be re-constructed, the challenge might lie with the comprehension of the story. Difficulty understanding the message to be reconstructed might have caused lower level learners to pay less attention to form than more proficient dyads, as they might not know what to talk about (Leeser, 2004; Swain & Lapkin, 2006). In contrast, the tasks used in the current study were open-ended tasks (though one is more open than the other) requiring learners to make their own meanings. That lower proficiency dyads discussed more LREs suggests they had noticed greater 'holes' or 'gaps' to fill or at least had more difficulty in finding resources to express their communicative intentions (Ellis et al., 2001a, 2001b) *in the face of the current tasks*. This shows that these lower proficiency dyads were more 'pushed' to carry out the same tasks than more proficient dyads, especially the HH. This indicates their 'problematicity' (Ellis et al., 2001a; Long, 2007; Williams, 2005) that reflects learners' real need to "talk about what they *need* to talk about" (Swain, 1998, p.73, original italics) in the context of making meanings, especially in communicative tasks.

An absence of a public performance in the prior studies could be another explanation for the discrepancy in the results. The incentive of the impending public performance could have catalysed the learners, especially the lower proficiency learners, given their more limited resources to prepare for the meanings they wanted to say to an audience. This pushed them to collaborate to resource their forthcoming co-performance. Research has shown that the nature of learner-learner interaction may influence the frequency of LREs (Dobao, 2012; Storch, 2002a, 2002b; Watanabe, 2008; Watanabe & Swain, 2007). In the current study, the few cases of less collaborative interactions that occurred belonged to the HH dyads, while the dialogues for the LL and the HL, were typically collaborative or with a clear expert/novice pattern respectively (see Storch, 2002a, 2002b). It is possible then that the more collaborative interaction in the HL and the LL dyads might have accounted for the higher occurrences of LREs from these groups. Williams (1999) argues that lower proficiency learners cannot attend to form to the same extent as more proficient learners because

there exists an “abyss”, not just a gap between their interlanguage and the target language (p.612). This might not necessarily be the case. It may not be the gap *per se*, but what triggers such a gap may be important. The findings show that when driven by meanings to convey to an audience (performance), and given an opportunity to prepare for that performance (rehearsal), low proficiency learners can attend to form even more frequently than high proficiency learners.

Regarding LRE resolutions, overall, the findings showed that learners were able to correctly resolve a large majority of their language problems. Similar results have been reported in other studies using dictogloss tasks (e.g., Alegría de la Colina & García Mayo, 2007; Kim & McDonough, 2008; Leaser, 2004; Malmqvist, 2005), text-editing tasks (Storch, 2007), information gap and ranking tasks (Newton, 2001, 2013) and a range of other communicative activities (McDonough & Sunitham, 2009). The success here, again, could be understood by ‘collaborative dialogue’ in the service of not only communication at hand (rehearsal), but also communication afterwards (performance) and by learner positive orientation to that performance (see Chapter 6). The learners predominantly sought assistance from each other rather than from the teacher. This contrasts with studies such as Williams (1999) in which requests for assistance were made of and resolved by the teacher, indicating a high level of learner activeness and autonomy in initiating, discussing and resolving language problems encountered. Some studies have pointed to learner autonomy as a factor in student-initiated focus on form (Loewen, 2006; Poole, 2005). Given the elite nature of the school, this is likely here since high motivation and autonomy arguably correlate.

Despite the overall success in resolving language problems encountered, lower proficiency dyads were less successful at resolving their language issues than higher proficiency dyads. This finding broadly corroborates those of previous studies (Kim & McDonough, 2008; Leaser, 2004; Watanabe & Swain, 2008). It is hardly surprising that the LL dyads did not resolve their language problems as well as the other more proficient dyads (HH, HL), given their lower linguistic resources. However, that the HL dyads were as able to correctly resolve LREs as the HH group is rather more interesting. The data showed that although dyad members assisted each other to solve language issues, in the HL dyads, the higher proficiency learner typically took the

expert role, providing solutions to language issues their lower proficiency peer encountered (see Episode 6), or responding to the peer's error (Episode 7).

Episode 6 (The problem-solving task-HL-11a)

Low: có thể ... có thể tạo điều kiện... ê tạo điều kiện là gì mi?

(*can ... can facilitate ... hey, how to say 'facilitate'?*)

High: facilitate

Low: tạo điều kiện (*facilitate*)

High: facilitate

Low: facilitate

Episode 7 (The problem-solving task-HL-11a)

Low: have you (.) have you (.) rằng hè (*how to say*) (.) have you done volunteer work?

High: have you EVER

Low: have you ever done volunteer work?

Indeed further analysis (Table 8.7) shows that of 81% ($M = .81$) of the LREs that the HL dyads resolved correctly, the higher proficiency peer resolved correctly more of them ($M = .64$) than the lower proficiency peer ($M = .17$) (whereas the latter initiated significantly more LREs, $M = 16.62$ vs. $M = 10.62$). It is the assistance of the more proficient learner in the HL dyad that levelled off the discrepancy in the overall proficiency between the HL and HH, accounting for the same level of success in finding correct LRE resolutions between the two dyad groups.

TABLE 8.7: Individual data for the HL dyads

	HL (8dyads)			H in HL (n=8)			L in HL (n=8)		
	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>
Correctly solved LREs	176	.81	.08	136	.64	.10	40	.17	.06
Frequency of LREs	218	27.25	5.90	85	10.62	3.02	133	16.62	4.74

The findings also show that learners sometimes incorrectly solved their language issues, which is consistent with previous research (Adams, 2007; Kowal & Swain, 1997; Pica, 1994; Zhao & Bitchener, 2007) finding that incorrect feedback did occur, though not frequently. The current study further shows that the proportions of incorrectly resolved LREs were not the same for all proficiency groups. They were more likely for the LL dyads (on average 2.5 times in every 10 LREs compared to 0.4 times out of 10 LREs for the HH and 1.3 times for the HL). All this provides useful information for the teacher as to how to manage feedback according to specific student pairings for task work (see Chapter 10).

The non-significant effect of proficiency on the incidence of unsolved LREs contrasts with findings from prior research which investigated the dictogloss task (Kim & McDonough, 2008; Leeson, 2004). For example, Leeson found that the lower proficiency dyads left 33.33% of their LREs unsolved when they reconstructed the passage of a dictogloss in pairs, a figure significantly higher than that of the HH (4%) and the HL (16%) in the same study. Similarly Kim and McDonough (2008) found that when two Korean as second language intermediate students worked together on a dictogloss task, they left 28% LREs unresolved. The LL group in the current study left on average only 9% of their LREs unresolved (compared to 5.6% and 5% for the HH and HL respectively). Once again, the overriding power of the public performance possibly drove the learners, regardless of proficiency to work out solutions to language issues encountered (though sometimes incorrectly) than leave them unsolved.

8.3.3 Rehearsal and the effects of linguistic focus of LREs on LRE resolutions

The results show that all dyad groups, irrespective of tasks, produced more lexical than grammatical LREs in task rehearsal. However, learners, especially the lower proficiency dyads, were more successful at resolving correctly grammatical than lexical issues.

The lexical predominance is due to the highly meaning-focussed and open-ended nature of the tasks used. Other research which investigated meaning-focussed tasks in teacher-learner interaction (Ellis et al., 2001a, 2001b; Loewen, 2004; Nassaji, 2010; Philp, Walter, & Basturkmen, 2010; Poole, 2005; Williams, 1999; Zhao & Bitchener, 2007) and in learner-learner interaction (Bitchener, 2004; Fuji & Mackey, 2009; Gass et

al., 2005; Hanaoka, 2007; Park, 2010) also found vocabulary predominant in focus. However, investigations into more form-focussed tasks such as dictogloss tasks or text-reconstruction tasks found more grammatical than lexical LREs (e.g., Alegría de la Colina & García Mayo, 2007; Kim & McDonough, 2008; Leeson, 2004; Malmqvist, 2005; Swain & Lapkin, 1998, 2001). It is clear that the nature of the tasks was an underlying factor. Consistent with Pica's (1997) findings, the learners in the current study did not pay attention to grammar as much as vocabulary because the tasks did not incline them to, not because they were not able to do so. The need to make meanings in open-ended meaning-focussed tasks inclined learners to pay more attention to lexis, because vocabulary carries most meaning (VanPatten, 1990, 1996). According to Skehan and Foster (2001), attention to vocabulary rather than grammar is not due to meaning and form competing for attention. Rather a meaning can still be conveyed efficiently without correct grammatical forms.

Of particular interest is that preference of attention to lexical items was consistent for all proficiency groups. This contrasts with other research which found more proficient learners not only produced more LREs, but also more grammatical LREs (e.g. Leeson, 2004). Importantly, the learners here chose to discuss LREs in order to problem-solve their meaning-making. Their LREs were therefore "learner-driven" or "learner-regulated" focus on form (Ortega, 2005, p.107), and addressed their own "problematicity" (Ellis, 2001; Long, 2007; Williams, 2005). Learners attended primarily to vocabulary, since it was the area of most 'problematicity' to them. Learners recognised the 'hole' between what they can say, and what they want to say (Swain, 1995, 2000, 2005). So lexical LREs represented learners' efforts to fill these holes to express intended meanings in the course of preparing, planning and rehearsing for the public task performance.

The nature of LREs in the data show how stretched and pushed learners were in mobilizing and searching for lexical items to convey their intended meanings. These LREs show a wide range of vocabulary being discussed, and stand in stark contrast to the "impoverished" and "minimalised" interaction that Seedhouse (1999), among others, have argued is typical of task-based interaction. The learners also reflected on each other's use of grammatical items, though to a smaller extent. The findings

contrast with doubts raised about the value of tasks in EFL contexts where students share an L1 and lack a need to communicate in the target language outside the classrooms (Carless, 2003, 2004, 2007; McDonough, 2004; Pham, 2007; see Adams & Newton, 2009; Butler, 2011 for recent reviews). Once again this shows that the quality of interaction is dependent not only on the task used, but also on the condition (e.g., rehearsal-performance) in which the task is carried out.

As regards LRE resolutions, one plausible explanation for why students were more successful with grammatical than lexical issues was that the grammatical problems learners discussed were easier to resolve than their lexical counterparts. These grammatical LREs typically involved familiar grammatical items that they had declarative knowledge of after all the years of learning. Typically students responded to their interlocutor's error or corrected themselves (Episode 8). In contrast, lexical problems could involve any target lexical items that learners wanted to use to express their intended meanings. The data showed that learners often asked each other for a wide range of vocabulary items that they were not always able to find (Episode 9).

Episode 8 (The debate task- HL-11f)

- S1: with my heart and er (.). Let's helping the world
S2: hah?
S1: let's helping the country
S2: let's help our country the world (.) let's help the country
S1: let's help our country (.) ah let đi với bare infinitive (*ah let plus bare infinitive verb*)

Episode 9 (The problem-solving task-LL-11c)

- S1: do cực khổ, cực khổ, cực khổ là chi hèo?
(*because of their strained life, strained life, how to say 'strained life'?*)
S2: vất vả đi
(*let's say 'hard life'*)
S1: vất vả là chi?
(*How to say 'hard life'?*)
S2: uhm

Therefore, in order to resolve grammatical problems, learners were more likely to have explicit grammar rules to resort to, whereas they did not have straightforward rules to weigh up lexical resolutions, particularly for problems that have multiple solutions (see Kormos, 2000). Episode 10 shows learners were trying to express the L1 meaning '*impossible*'. They came up with a few solutions such as '*incredible*', '*fiction*', '*fantastic*', '*romantic*' and finally settled down on '*unacceptable*', which was an incorrect one. This shows how difficult it is for learners to reach a definitive solution for lexical problems.

Episode 10 (The problem-solving task-HH-11f)

- S2: ví dụ tau nói (*for example I'll say*) have you ever er had 500 million?
- S1: không tưởng, incredible không thể tin được
(*impossible, incredible means unbelievable*)
- S2: incredible?
- S1: incredible là không thể tin được, không tưởng là chi hè?
(*incredible means unbelievable, how to say impossible?*)
- S2: fiction là viễn tưởng [self-translating]
- S1: điều không tưởng là chi hèo (*how to say impossible?*)? Fantastic (.) điều không tưởng er điều viễn vông (*impossible, er not practical*) (.) romantic [laugh] là viễn vông xa vời (*is impossible*)
- S2: unacceptable là không thể chấp nhận được [self-translating]
- S1: unacceptable. Rồi chưa? Xong ba lý do chưa? (*Done? Have we got three reasons?*)

Furthermore, knowing a word is more than knowing the meaning of that word, but knowing its collocations, knowing how to use it grammatically and syntactically in meaningful contexts (Nation, 2001), making it more difficult to resolve lexical problems during meaningful communication. Episode 11 is an illustrative example.

Episode 11 (The debate task-LL-11b)

- S1: the poor have (.) don't have enough money to eat
S2: họ làm từ thiện bằng cách nào là how do they charity?
(*'How do they do charity' is 'how do they charity'?*)
S1: how DO THEY DO VOLUNTEER![Emphasizing]
S2: charity là cũng làm từ thiện rồi
(*charity is also a verb already*)
S1: rồi (OK)

Here in response to S1's idea that *'poor people don't have money to eat'*, S2 wants to say *'how do they do charity?'* She questions her English expression *'how do they charity?'* by asking her interlocutor, who suggests quite strongly *'how do they do volunteer'* (she apparently wants to mean *'how do they do volunteer work?'*). S1, at this point defends her original formulation by incorrectly explaining that *'charity'* by itself is a verb. S2 appears to be led by S1 by now and seems to agree (*'OK'*). This example shows that encoding the meaning of *'how do they do charity?'* involves more than knowing the word *'charity'*, but also its parts of speech, and its collocation- *'do charity'* and using it in a sentence. This may explain why all proficiency groups abandoned a higher proportion of lexical than grammatical LREs unsolved and why learners, particularly low proficiency learners, incorrectly resolved a higher proportion of their lexical than grammatical LREs.

Given the burden involved in resolving lexical issues, for the HH and HL dyads, when the L2 resource was higher, they were able to correctly resolve both lexical and grammatical LREs equally well. For the LL dyads, when their resource was low, they were less successful at resolving lexical than grammatical problems. The finding is consistent with prior research which investigated low proficiency learners. For example, McDonough and Sunitham (2009) found low proficiency students at an EFL Thai university solved correctly more grammatical LREs and incorrectly resolved or left unresolved more lexical LREs, though these differences were not significant. This could be due to the small sample data. Iwashita (2001), through examining tasks from a different angle (negotiation of meaning and modified output) argued that syntactic items were within the learners' knowledge to modify while they lacked vocabulary to

generate lexical modifications. The same argument Iwashita mounted was that morphosyntactic knowledge is more ‘manageable’. The current findings further suggest that the effects of linguistic focus may be mediated by learner proficiency.

Another interpretation as to why lower proficiency dyads were better at correctly resolving grammatical than lexical problems could be related to their greater orientation towards grammatical than lexical accuracy in the public performance. The following excerpt from a LL dyad seems to corroborate this explanation.

- S1: Đừng lo, chắc chi bọn nớ biết từ ni. Chi có cô biết thôi, ah cô nớ nữa. [laugh]
 (Don't worry, our classmates might not know this word, except the teacher, ah and that teacher. [the researcher][laugh]
- S2: Ừ, cố gắng đừng sai ngữ pháp nhiều quá là được.
 (Yes, try not to make so many grammatical mistakes and this is fine.)
- S1: Ừ. [laugh]
 (Yes.)

Under the pressure of an imminent public performance, coupled with their low resource, low proficiency dyads were more aware of their own or their peer's grammatical mistakes (as evidenced by more grammatical LREs than the other groups), leading to greater effort to respond to and resolve them.

8.4 Summary

To recap, the findings have shown tasks and proficiency affected not only the occurrences of LREs, but also how they were resolved in task rehearsal. However, regardless of tasks and proficiency, learners encountered more lexical than grammatical problems. Yet the LL dyads were better at correctly resolving grammatical than lexical LREs, whereas the HH and HL dyads were able to correctly resolve both lexical and grammatical issues equally well. Above all, learners were ‘pushed’ to find correct forms to express the meanings to be communicated to the class audience (performance). If LREs represent ‘occasions for learning’ (Swain & Lapkin, 1998) or ‘a source of learning’ (Swain, 2001) in collaborative task rehearsal, I am now looking for evidence of this learning occurring, via uptake in the performance itself and how this

uptake is affected by task type, proficiency and the linguistic focus of LREs. This will be the focus of Chapter 9.

Chapter 9 THE PHASE 2 STUDY: UPTAKE IN TASK PERFORMANCE

9.1 Introduction

Building on the results on LREs in task rehearsal reported in Chapter 8, this chapter focuses on the extent to which language items attended to in LREs in task rehearsal were taken up in task performance. It addresses the following research questions (RQ).

- RQ1. To what extent did learning opportunities operationalised as LREs in rehearsal lead to uptake in performance? Did task type and proficiency impact on the level of uptake?
- RQ2. Did task type and proficiency affect different types of uptake in task performance?
- RQ3. Was there a relationship between LRE resolutions in rehearsal and uptake in performance?
- RQ4. Did uptake in task performance differ by linguistic focus?

I will first report on the findings related to each of the research questions and then discuss some main results and conclude with a summary.

9.2 Findings

9.2.1 To what extent did learning opportunities operationalised as LREs in rehearsal lead to uptake in performance? Did task type and proficiency impact on the level of uptake?

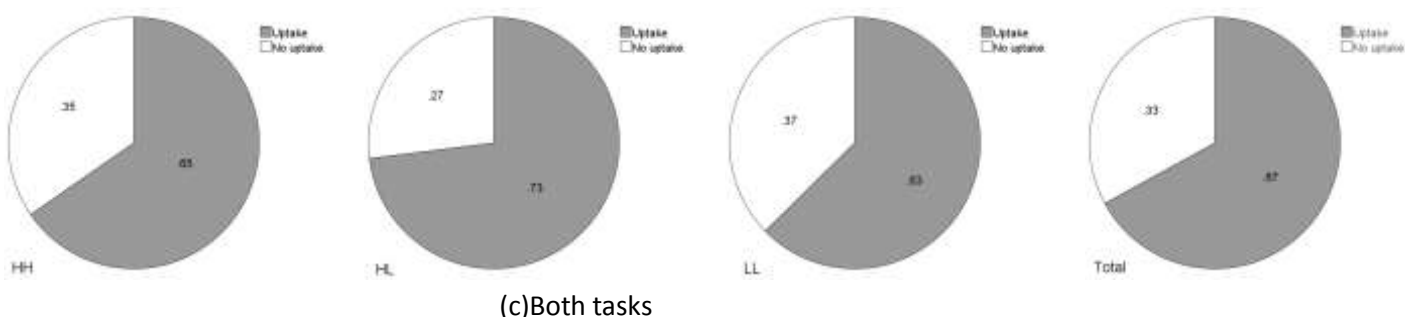
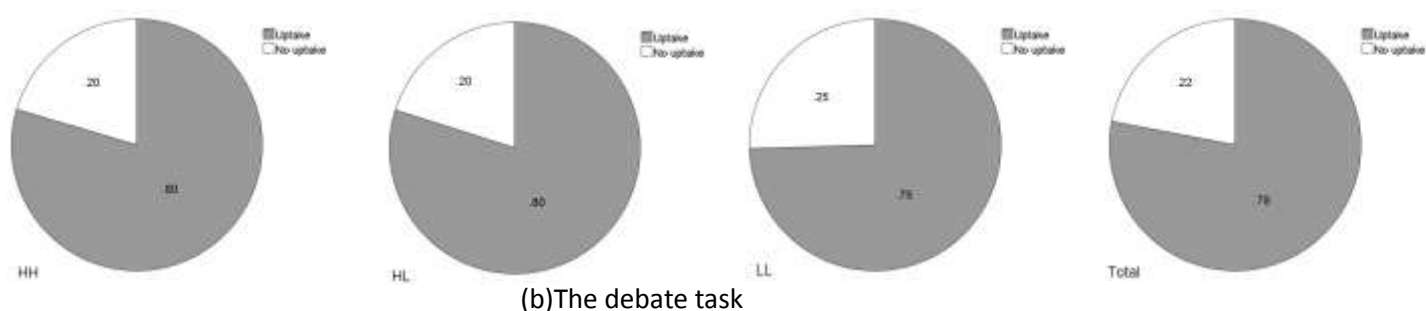
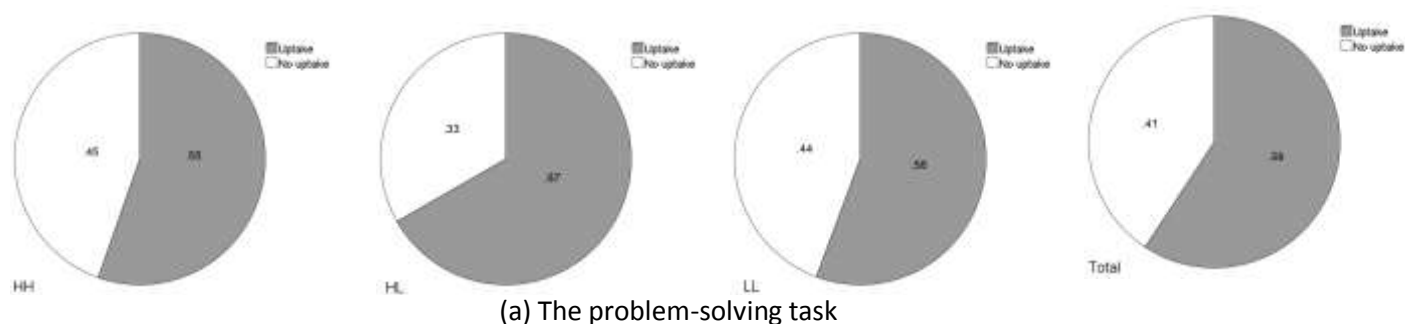
As described in Chapter 7, instances of language items focussed on in LREs in task rehearsal that *were not used at all* in any way in task performance (regardless of their resolutions) were coded as *no uptake*. If the language item targeted in a certain LRE, (be it correctly resolved, incorrectly resolved or unsolved) in task rehearsal was used in task performance, it was coded as *uptake*. Table 9.1 shows the total uptake across tasks and proficiency groups which is further displayed in Figure 9.1.

TABLE 9.1: Total uptake across tasks and proficiency groups

		HH (n=8 dyads)			HL (n=8 dyads)			LL (n=8 dyads)			Total		
		<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>
Uptake	T1	53	.55	.11	79	.67	.10	86	.56	.06	218	.59	.10
	T2	58	.80	.13	79	.80	.15	73	.75	.18	210	.78	.15
	<i>Total</i>	<i>111</i>	<i>.65</i>	<i>.07</i>	<i>158</i>	<i>.73</i>	<i>.07</i>	<i>159</i>	<i>.63</i>	<i>.07</i>	<i>428</i>	<i>.67</i>	<i>.08</i>
No uptake	T1	46	.45	.11	38	.33	.10	67	.44	.06	151	.41	.10
	T2	16	.20	.13	22	.20	.15	31	.25	.18	69	.22	.15
	<i>Total</i>	<i>62</i>	<i>.35</i>	<i>.07</i>	<i>60</i>	<i>.27</i>	<i>.07</i>	<i>98</i>	<i>.37</i>	<i>.07</i>	<i>220</i>	<i>.33</i>	<i>.08</i>
Frequency of LREs	T1	99	12.38	4.34	117	14.63	4.03	153	19.13	6.96	369	15.38	5.80
	T2	74	9.25	4.13	101	12.63	4.17	104	13.00	4.34	279	11.63	4.38
	<i>Total</i>	<i>173</i>	<i>21.63</i>	<i>8.16</i>	<i>218</i>	<i>27.25</i>	<i>5.90</i>	<i>257</i>	<i>32.13</i>	<i>9.13</i>	<i>648</i>	<i>27.00</i>	<i>8.68</i>

Note. T1= The problem-solving task; T2= The debate task

FIGURE 9.1: Total uptake by task and proficiency



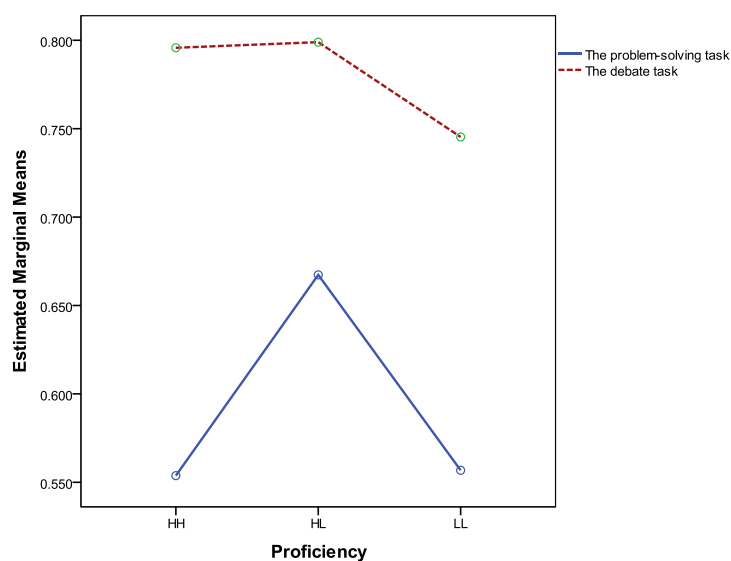
As Table 9.1 and Figure 9.1 show, the level of uptake was high, on average at 67 % in total. The debate task seemed to yield more uptake than the problem-solving task, and all proficiency groups seemed to have similar rates of uptake for each individual task and for the two tasks combined. In order to test the significance of these trends, an RM ANOVA, with one within-subject variable, task type (two levels: the problem-solving task and the debate task), and one between-subject variable, proficiency (three levels: HH, HL, LL) was carried out and the results are summarised in Table 9.2.

TABLE 9.2: Summary of the RM ANOVA results for the total uptake in task performance

Source		<i>F</i>	<i>p</i>	η_p^2	Observed <i>power</i>
<i>Within-subjects</i>	Task	24.601	.000	.54	1.00
	Task*	.712	.502	.06	.15
	proficiency				
<i>Between-subjects</i>	Proficiency	1.982	.163	.16	.36

The results showed a significant main effect of task, $F(1, 21) = 24.601$, $p < .001$, with a very large effect size ($\eta_p^2 = .54$) and high power (100%), indicating that, proficiency aside, learners achieved a significantly higher proportion of uptake when performing the debate task ($M = .78$, $SD = .15$) than the problem-solving task ($M = .59$, $SD = .10$). The interaction effect between task type and proficiency was not statistically significant, $F(2, 21) = .712$, $p = .502$, indicating that all groups had a higher level of uptake on the performance of the debate task (see Figure 9.2). The results also showed no significant main effect of proficiency, $F(2, 21) = 1.982$, $p = .163$. The mean proportions of total uptake for the groups were: for the HH, $M = .65$ ($SD = .07$), for the HL, $M = .73$ ($SD = .07$) and for the LL, $M = .63$ ($SD = .07$). Put simply, whether the dyads were two higher proficiency learners (HH), or two lower proficiency learners (LL), or mixed level dyads (HL), they all achieved similar high rates of total uptake.

FIGURE 9.2: Total uptake across tasks and proficiency groups



In summary, irrespective of proficiency, a majority of language items attended to in LREs in rehearsal were used in performance. However, the uptake rate was significantly higher when learners performed the debate task than the problem-solving task.

9.2.2 Did task type and proficiency affect different types of uptake in task performance?

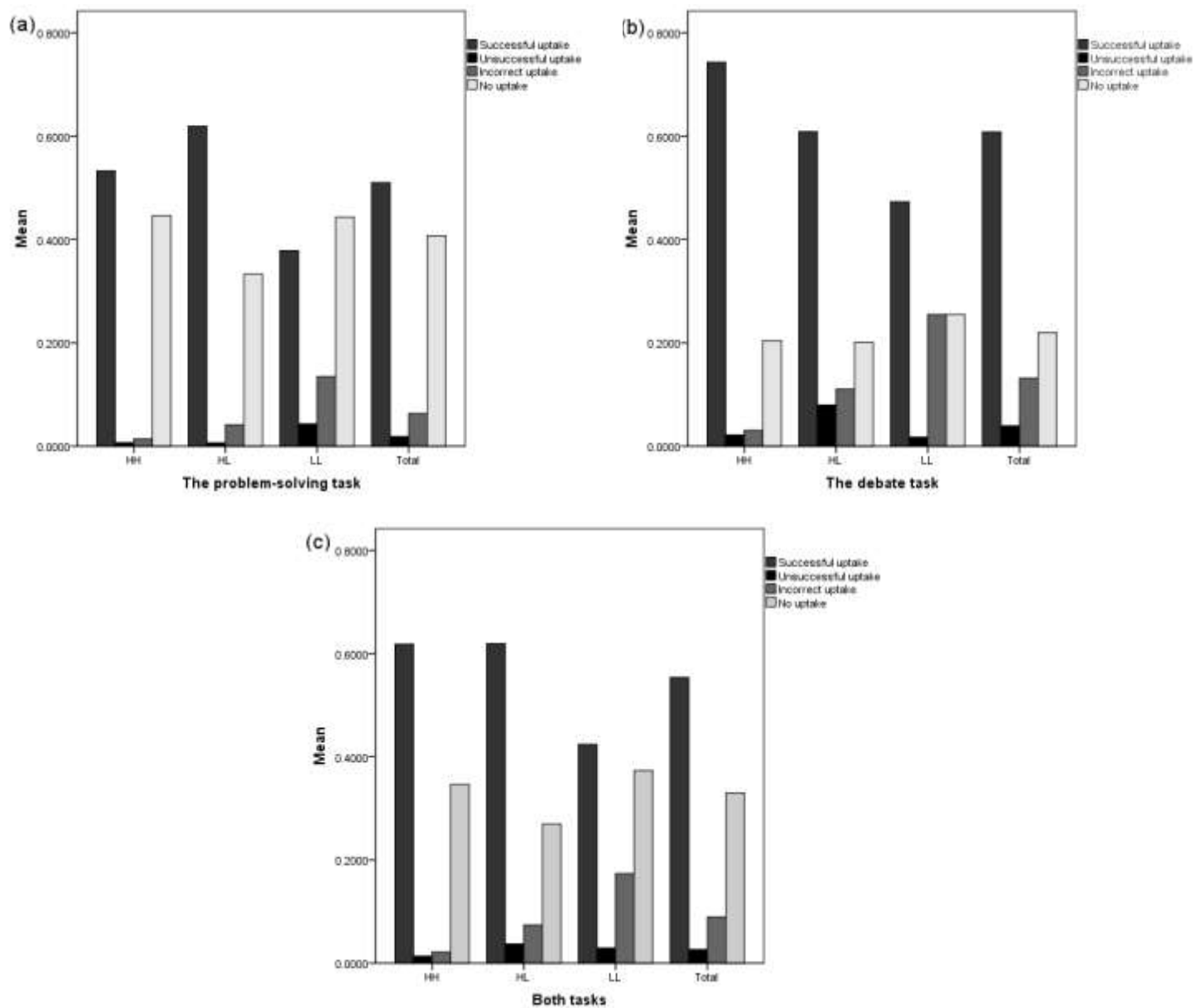
To address this question, uptake of the targets of LREs in rehearsal was coded as either successful, unsuccessful, or incorrect uptake (see Chapter 7 for the detailed operational definitions of these types of uptake). Table 9.3 shows the data for these uptake types across tasks and proficiency groups. The means of these uptake types are displayed in Figure 9.3.

TABLE 9.3: Uptake types in task performance by task and proficiency

		HH(n=8 dyads)			HL(n=8 dyads)			LL(n=8 dyads)			Total		
Uptake		<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>	<i>Sum</i>	<i>M</i>	<i>SD</i>
Successful uptake	T1	50	.53	.13	73	.62	.11	59	.38	.09	182	.51	.15
	T2	53	.74	.16	60	.61	.16	46	.47	.17	159	.61	.19
	<i>Total</i>	<i>103</i>	<i>.62</i>	<i>.10</i>	<i>133</i>	<i>.62</i>	<i>.08</i>	<i>105</i>	<i>.42</i>	<i>.09</i>	<i>341</i>	<i>.55</i>	<i>.13</i>
Un-successful uptake	T1	1	.01	.02	1	.01	.02	6	.04	.04	8	.02	.03
	T2	2	.02	.04	7	.08	.09	2	.02	.03	11	.04	.06
	<i>Total</i>	<i>3</i>	<i>.01</i>	<i>.03</i>	<i>8</i>	<i>.04</i>	<i>.04</i>	<i>8</i>	<i>.03</i>	<i>.02</i>	<i>19</i>	<i>.03</i>	<i>.03</i>
Incorrect uptake	T1	2	.01	.03	5	.04	.06	21	.13	.05	28	.06	.07
	T2	3	.03	.04	12	.11	.08	25	.25	.19	40	.13	.15
	<i>Total</i>	<i>5</i>	<i>.02</i>	<i>.03</i>	<i>17</i>	<i>.07</i>	<i>.06</i>	<i>46</i>	<i>.17</i>	<i>.05</i>	<i>68</i>	<i>.09</i>	<i>.08</i>
No uptake	T1	46	.45	.11	38	.33	.10	67	.44	.06	151	.41	.10
	T2	16	.20	.13	22	.20	.15	31	.25	.18	69	.22	.15
	<i>Total</i>	<i>62</i>	<i>.35</i>	<i>.07</i>	<i>60</i>	<i>.27</i>	<i>.07</i>	<i>98</i>	<i>.37</i>	<i>.07</i>	<i>220</i>	<i>.33</i>	<i>.08</i>
Frequency of LREs	T1	99	12.38	4.34	117	14.63	4.03	153	19.13	6.96	369	15.38	5.80
	T2	74	9.25	4.13	101	12.63	4.17	104	13.00	4.34	279	11.63	4.38
	<i>Total</i>	<i>173</i>	<i>21.63</i>	<i>8.16</i>	<i>218</i>	<i>27.25</i>	<i>5.90</i>	<i>257</i>	<i>32.13</i>	<i>9.13</i>	<i>648</i>	<i>27.00</i>	<i>8.68</i>

Note. T1= The problem-solving task; T2= The debate task

FIGURE 9.3: Uptake types across tasks and proficiency groups



In order to test the effects of task type and proficiency on each type of uptake, separate RM ANOVAs were performed, with task type as a within-subject variable (with two levels: the problem-solving and the debate task) and proficiency (with three levels: HH, HL, LL) as a between-subject variable. The RM ANOVA results are provided in Table 9.4.

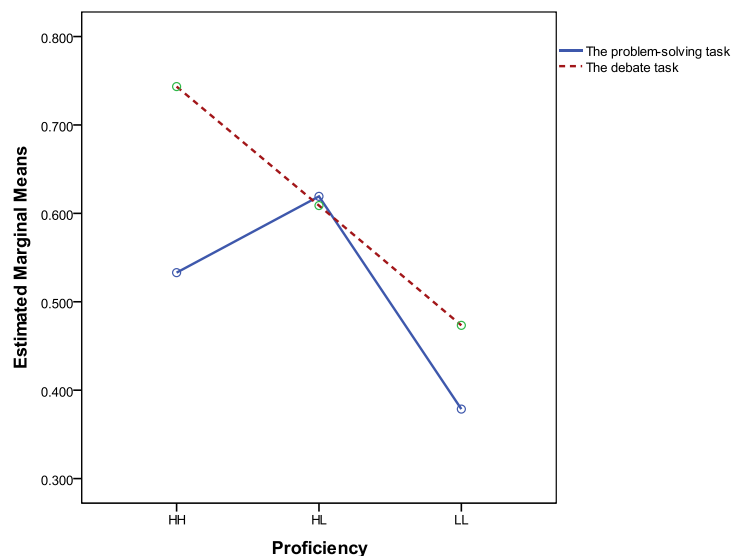
TABLE 9.4: Summary of the RM ANOVA results for uptake types in task performance

Source		Measure	<i>F</i>	<i>p</i>	η_p^2	Observed <i>power</i>
<i>Within-subjects</i>	Task	Successful uptake	6.576	.018	.24	.69
		Unsuccessful uptake	3.036	.096	.13	.38
		Incorrect uptake	6.355	.020	.23	.67
	Task type* Proficiency	Successful uptake	2.771	.086	.21	.49
		Unsuccessful uptake	5.891	.009	.36	.82
		Incorrect uptake	1.199	.321	.10	.23
<i>Between-subjects</i>	Proficiency	Successful uptake	9.878	.001	.49	.97
		Unsuccessful uptake	1.387	.272	.12	.27
		Incorrect uptake	15.179	.000	.59	1.00

Regarding successful uptake, the results showed a significant main effect of task, $F(1, 21) = 6.576$, $p = .018$, $\eta_p^2 = .24$, suggesting that proficiency aside, learners attained significantly more successful uptake in the debate task ($M = .61$, $SD = .19$) than in the problem-solving task ($M = .51$, $SD = .15$). The main effect of proficiency was also statistically significant, $F(2, 21) = 9.878$, $p = .001$, with a large effect size ($\eta_p^2 = .49$) and high power of 97%. The mean proportions of successful uptake for the groups were: for the HH, $M = .62$ ($SD = .10$), for the HL, $M = .62$ ($SD = .08$) and for the LL, $M = .42$ ($SD = .09$). Tukey post-hoc analysis showed that the LL group had a significantly lower level of successful uptake than the HH ($p = .002$) and the HL ($p = .005$), but there was no significant difference between the HH and HL ($p = .890$). The interaction effect between task type and proficiency was not statistically significant, $F(2, 21) = 2.771$, $p = .086$. However, it was approaching significance ($p = .086$) that accounted for 21% ($\eta_p^2 = .21$) of the variance and this warrants discussion (Larson-Hall, 2010). As shown in

Figure 9.4, for the HL group, the successful uptake rates on the two tasks were almost similar, whereas the level of successful uptake was greater in favour of the debate task for the remaining groups. Further paired sample t -tests showed no significant difference in the levels of successful uptake between the two tasks for the HL dyads, $t(7) = .141, p = .892, r = .06$, and for the LL dyads, $t(7) = -1.495, p = .178, r = .49$, (though the effect size was large, $r = .49$). The mean proportions of successful uptake in the problem-solving task and the debate task respectively were: for the LL, $M = .38$ ($SD = .09$) and $M = .47$ ($SD = .17$); for the HL, $M = .62$ ($SD = .11$) and $M = .61$ ($SD = .16$). However, the HH group achieved a significantly higher rate of successful uptake on the debate task ($M = .74, SD = .16$) than the problem-solving task ($M = .53, SD = .13$), $t(7) = -3.426, p = .011, r = .79$.

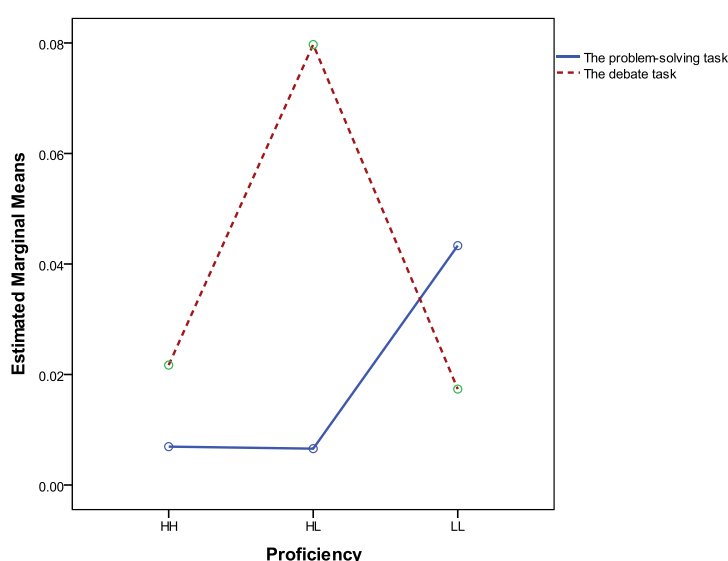
FIGURE 9.4: Successful uptake across tasks and proficiency groups



For unsuccessful uptake, the results showed no significant main effect of task, $F(1, 21) = 3.036, p = .096$, though it was approaching significance. The mean proportions were: $M = .02$ ($SD = .03$) for the problem-solving task and $M = .04$ ($SD = .06$) for the debate task. The main effect of proficiency was not significant, $F(2, 21) = 1.387, p = .272$. The mean proportions of unsuccessful uptake for the groups were: for the HH, $M = .01$ ($SD = .03$), for the HL, $M = .04$ ($SD = .04$) and for the LL, $M = .03$ ($SD = .02$). However, the interaction effect between task and proficiency was statistically significant, $F(2, 21) =$

5.891, $p = .009$, $\eta_p^2 = .36$, indicating that levels of unsuccessful uptake differed, depending on the task (see Figure 9.5). The HL and the HH groups produced more unsuccessful uptake when doing the debate task whereas the LL group did so on the problem-solving task. Given the very low frequency of unsuccessful uptake (from 1-7 instances), these results should be read with caution, as small numbers may mislead the proportions. Above all, the results showed that the learners infrequently used correct resolutions from rehearsal incorrectly in task performance.

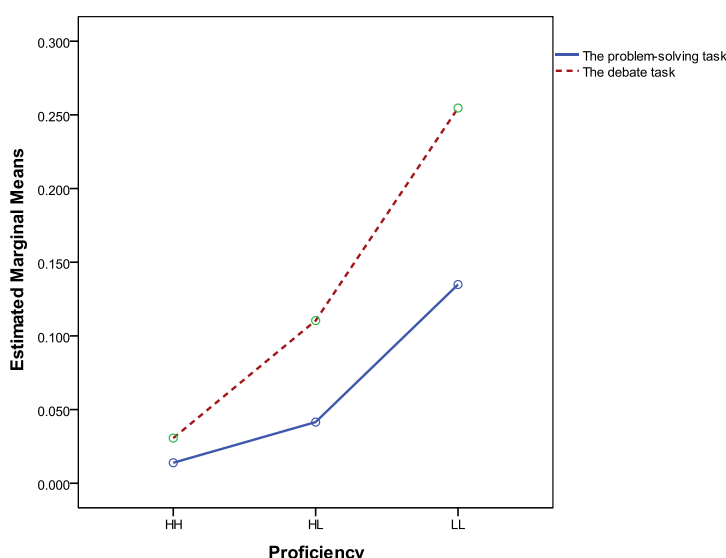
FIGURE 9.5: Unsuccessful uptake across tasks and proficiency groups



Regarding how task type and proficiency affected incorrect uptake, the results showed a statistically significant effect of task, $F(1, 21) = 6.355$, $p = .020$, $\eta_p^2 = .23$. The learners had a higher proportion of incorrect uptake on the performance of the debate task ($M = .25$, $SD = .19$) than the problem-solving task ($M = .13$, $SD = .05$). This trend was consistent for all proficiency groups, with no significant interaction between task and proficiency, $F(2, 21) = 1.199$, $p = .321$ (see Figure 9.6). The main effect of proficiency on incorrect uptake was also statistically significant, $F(2, 21) = 15.179$, $p < .001$, with a very large effect size ($\eta_p^2 = .59$), and high power of 100%. The mean proportions of incorrect uptake for the groups were: for the HH, $M = .02$ ($SD = .03$), or the HL, $M = .07$ ($SD = .06$) and for the LL, $M = .17$ ($SD = .05$). Tukey post-hoc analysis showed that the LL dyads had a statistically higher level of incorrect uptake than the HH ($p < .001$), and the

HL ($p = .004$). However, the difference between the HH and the HL was not statistically significant ($p = .238$).

FIGURE 9.6: Incorrect uptake across tasks and proficiency groups



In short, the results showed strong effects of tasks and proficiency on the quality of uptake in task performance. Dyads, especially the HH and LL, achieved a higher level of successful uptake in the performance of the debate task than the problem-solving task, but at the same time all proficiency groups had more incorrect uptake in the performance of the debate task. Lower proficiency dyads achieved a lower level of successful uptake than higher proficiency dyads, but were more likely to have incorrect uptake.

9.2.3 What was the relationship between LRE resolutions in task rehearsal and uptake in task performance?

This question asked how success at resolving language problems in task rehearsal was related to uptake in task performance. A series of correlation tests were conducted for each of the two tasks (the problem-solving task and the debate task). The data for correctly-solved LREs and successful uptake was normally distributed, so Pearson correlation tests (r) were used. Non-parametric correlation tests (Spearman's rho) (r_s) were conducted for the relationship between incorrectly solved LREs and incorrect uptake and between unsolved LREs and no uptake because these data showed non-

normal distribution. The results showed a positive relationship between LRE resolutions and the quality of uptake, irrespective of tasks (Table 9.5). For example, correct LRE resolutions in rehearsal were positively correlated with successful uptake in task performance: $r = .82, p < .001$ and $r = .81, p < .001$ for the problem-solving task and the debate task respectively. Similarly there was a positive association between incorrectly solved LREs and incorrect uptake for both the problem-solving task ($r_s = .89, p < .001$) and the debate task ($r_s = .87, p < .001$). Likewise, when learners left their LREs unsolved, these unsolved LREs had a positive correlation with no uptake: for the former task ($r_s = .51, p = .012$) and for the latter task ($r_s = .57, p = .004$). All these relationships were very strong ($r/r_s > .50$).

TABLE 9.5: Summary of correlation results for LRE resolutions by task

	1 st variable	2 nd variable	Correlation	<i>p</i>
The problem-solving task	Correctly solved LREs	Successful uptake	.82	.000
	Incorrectly solved LREs	Incorrect uptake	.89	.000
	Unsolved LREs	No uptake	.51	.012
The debate task	Correctly solved LREs	Successful uptake	.81	.000
	Incorrectly solved LREs	Incorrect uptake	.87	.000
	Unsolved LREs	No uptake	.57	.004

The data shows that unsuccessful uptake was infrequent with 19 out of the 495 correctly solved LREs (3.84%) being used incorrectly (unsuccessfully) in task performance (see Table 9.6). This further confirms that once the learners correctly resolved their language problems in the rehearsal, a majority of these correct language resolutions were adopted in task performance³⁴ (see Table 9.7): 61% for the problem-solving task and 76% for the debate task. For the total data, 67% of the correct language items attended to in LREs in rehearsal was transferred successfully to task performance: 61% for the LL dyads, 66% for the HH, and 74% for the HL. Chi-squared tests were conducted to see which proficiency group was more likely to use language resolutions correctly in task performance. The results showed the HL dyads were more

³⁴ The number of incorrectly solved LREs led to correct use in performance was minor (8 instances) (see 7.7.5.3).

likely to retain correct language resolutions in task performance than the LL dyads, $\chi^2(1) = 7.057, p = .010$. However, there were no significant differences between the LL and HH, $\chi^2(1) = .801, p = .416$, and the HL and HH, $\chi^2(1) = 2.942, p = .092$, indicating that once the lower proficiency dyads (LL) were able to correctly solve their language problems, they were able to use them in their performance equally as well as higher proficiency dyads (HH).

TABLE 9.6: Percentages of unsuccessful uptake out of correctly solved LREs

	Proficiency group	Total LREs	Correctly-solved LREs	Unsuccessful uptake per correctly solved LREs	
				n	%
The problem-solving task	HH	99	88	1	1.14
	HL	117	102	1	1.00
	LL	153	100	6	6.00
	<i>Total</i>	369	290	8	2.76
The debate task	HH	74	67	2	3.00
	HL	101	74	7	9.46
	LL	104	64	2	3.13
	<i>Total</i>	279	205	11	5.37
Both tasks	HH	173	155	3	1.94
	HL	218	176	8	4.55
	LL	257	164	8	4.88
	<i>Total</i>	648	495	19	3.84

TABLE 9.7: Percentages of successful uptake out of correctly resolved LREs

	Proficiency group	Total LREs	Correctly-solved LREs	Successful uptake per correctly solved LREs	
				N	%
The problem-solving task	HH	99	88	49	55.68
	HL	117	102	72	70.58
	LL	153	100	56	56.00
	<i>Total</i>	<i>369</i>	<i>290</i>	<i>177</i>	<i>61.03</i>
The debate task	HH	74	67	53	79.10
	HL	101	74	59	79.72
	LL	104	64	44	68.75
	<i>Total</i>	<i>279</i>	<i>205</i>	<i>156</i>	<i>76.09</i>
Both tasks	HH	173	155	102	65.80
	HL	218	176	131	74.43
	LL	257	164	100	60.97
	<i>Total</i>	<i>648</i>	<i>495</i>	<i>333</i>	<i>67.27</i>

Although incorrectly solved LREs in this study made up 16% (105/648 LREs) (see Chapter 8), of these, 65% (68/105) led to incorrect use in task performance: 62.50% (5/8) for the HH, 56.67% (17/30) for the HL, and 68.66% (46/67) for the LL (see Table 9.8). The Chi-squared results revealed no significant difference between the three proficiency groups, $\chi^2(2) = 1.325$, $p = .516$, suggesting that once students came up with incorrect resolutions to their language problems in rehearsal, irrespective of their proficiency, they used a large majority of them incorrectly in task performance.

None of the unsolved LREs led to uptake for both tasks and all dyads.

TABLE 9.8: Percentages of incorrect uptake out of incorrectly solved LREs

	Proficiency group	Total LREs	Incorrectly-solved LREs	Incorrect uptake per incorrectly solved LREs	
				N	%
The problem-solving task	HH	99	4	2	50.00
	HL	117	8	5	62.50
	LL	153	31	21	67.74
	<i>Total</i>	369	43	28	65.11
The debate task	HH	74	4	3	75.00
	HL	101	22	12	54.55
	LL	104	36	25	69.44
	<i>Total</i>	279	62	40	64.52
Total	HH	173	8	5	62.50
	HL	218	30	17	56.67
	LL	257	67	46	68.66
	<i>Total</i>	648	105	68	64.76

In brief, the ways learners resolved their LREs in task rehearsal (correctly resolved, incorrectly resolved and unresolved) were closely associated with the quality of uptake (successful, unsuccessful and incorrect) in task performance.

9.2.4 Did uptake in task performance differ by linguistic focus?

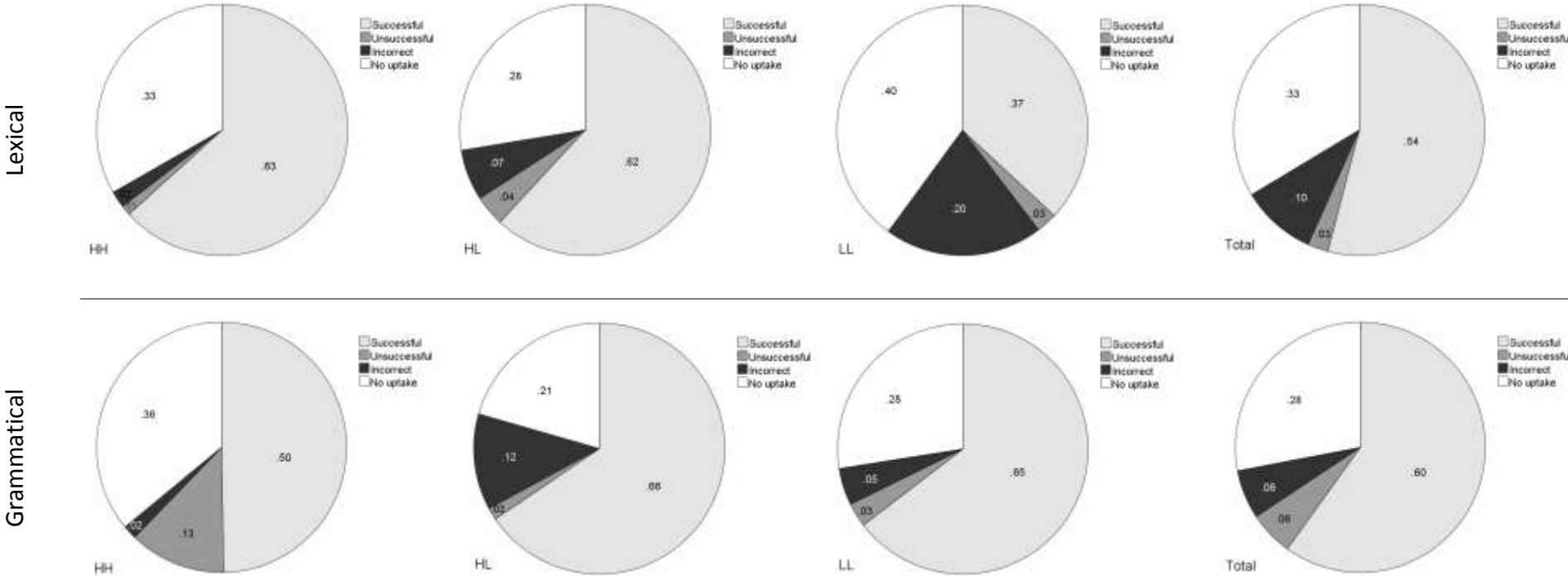
This question is concerned with whether lexical or grammatical LREs in rehearsal were more likely to be taken up in task performance. Due to the lower frequency of grammatical LREs relative to lexical LREs, the data from the two tasks were combined in order to assure adequate data for reliable statistical analyses. For each type of uptake, the proportion of lexical or grammatical uptake was calculated out of the total lexical or grammatical LREs. Table 9.9 shows the data on uptake types across linguistic foci and proficiency groups which are further displayed in Figure 9.7.

TABLE 9.9: Uptake types by linguistic focus across proficiency groups

	Successful		Unsuccessful		Incorrect		No uptake		Frequency of LREs	
	<i>L</i>	<i>G</i>	<i>L</i>	<i>G</i>	<i>L</i>	<i>G</i>	<i>L</i>	<i>G</i>	<i>L</i>	<i>G</i>
HH (n= 8 dyads)										
<i>Sum</i>	84	19	2	1	4	1	50	12	140	33
<i>M</i>	.63	.50	.01	.13	.02	.02	.33	.36		
<i>SD</i>	.18	.36	.03	.35	.03	.05	.14	.33		
HL (n= 8 dyads)										
<i>Sum</i>	110	23	6	2	12	5	52	8	180	38
<i>M</i>	.62	.66	.04	.02	.07	.12	.28	.21		
<i>SD</i>	.09	.22	.04	.04	.06	.17	.07	.20		
LL (n= 8 dyads)										
<i>Sum</i>	72	33	6	2	43	3	84	14	205	52
<i>M</i>	.37	.65	.03	.03	.20	.05	.40	.28		
<i>SD</i>	.11	.13	.02	.06	.07	.07	.09	.15		
Total										
<i>Sum</i>	266	75	14	5	59	9	186	34	525	123
<i>M</i>	.54	.60	.03	.06	.10	.06	.34	.28		
<i>SD</i>	.18	.26	.03	.21	.10	.12	.11	.24		

Note. L= Lexical; G= Grammatical

FIGURE 9.7: Uptake by linguistic focus across proficiency groups



In order to see whether linguistic focus and proficiency affect uptake, separate RM ANOVAs were performed on the data, with linguistic focus as one-within subject variable, and proficiency as one between-subject variable. The RM ANOVA results are summarised in Table 9.10.

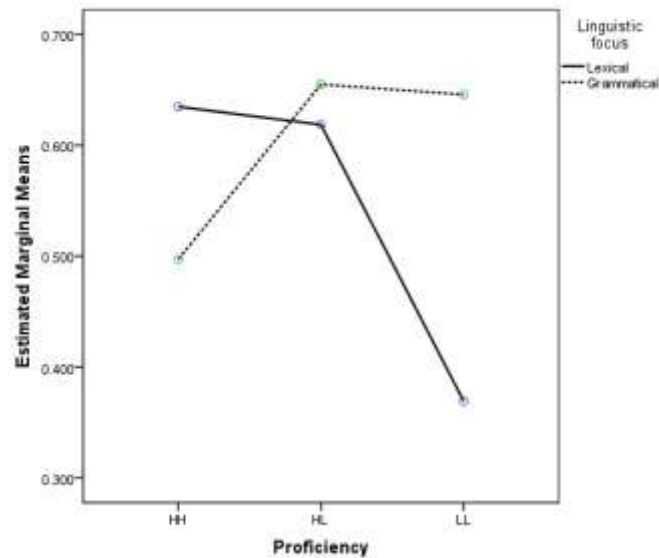
TABLE 9.10: Summary of the RM ANOVA results for uptake types across linguistic focus and proficiency

Source		Measure	<i>F</i>	<i>p</i>	η_p^2	Observed power
<i>Within-subjects</i>	Linguistic focus	Successful uptake	.933	.345	.04	.15
		Unsuccessful uptake	.218	.646	.01	.07
		Incorrect uptake	1.803	.194	.08	.25
		No uptake	1.012	.326	.05	.16
	Linguistic focus*	Successful uptake	3.937	.035	.27	.64
		Unsuccessful uptake	.815	.456	.07	.17
	Proficiency	Incorrect uptake	6.370	.007	.38	.85
		No uptake	.618	.548	.06	.14
<i>Between-subjects</i>	Proficiency	Successful uptake	1.735	.201	.14	.32
		Unsuccessful uptake	.368	.696	.03	.10
		Incorrect uptake	5.669	.011	.35	.81
		No uptake	1.813	.188	.15	.34

On the question of whether lexical or grammatical LREs were more likely to lead to successful uptake in task performance, the results showed that the interaction between linguistic focus and proficiency was significant, $F(2,21) = 3.937$, $p = .035$, $\eta_p^2 = .27$, indicating that the effect of linguistic focus on the likelihood of successful uptake was mediated by proficiency (see Figure 9.8). The LL dyads took up a significantly higher proportion of grammatical items ($M = .65$, $SD = .13$) successfully in task performance than lexical items ($M = .37$, $SD = .11$), $t(7) = -4.572$, $p = .003$, $r = .87$. The HL dyads also took up a higher proportion of grammatical items ($M = .66$, $SD = .22$) than lexical items ($M = .62$, $SD = .09$), but the difference was not significant, $t(7) = -.497$, $p = .635$, $r = .18$. In contrast, the HH dyads had a higher proportion of lexical ($M = .64$,

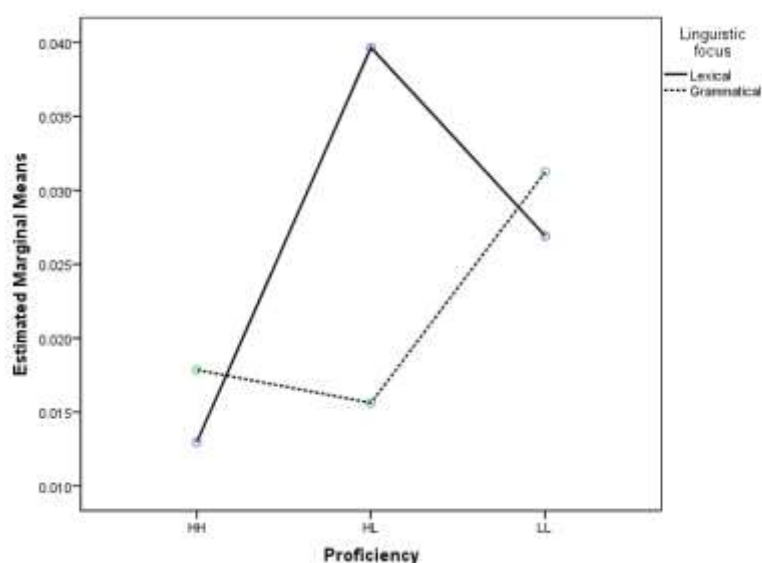
$SD=.18$) than grammatical items ($M= .50, SD =.36$) carried successfully to task performance, although again the difference was not statistically significant, $t(7) = .890$, $p = .403$, $r=.32$.

FIGURE 9.8: Successful uptake by linguistic focus across proficiency groups



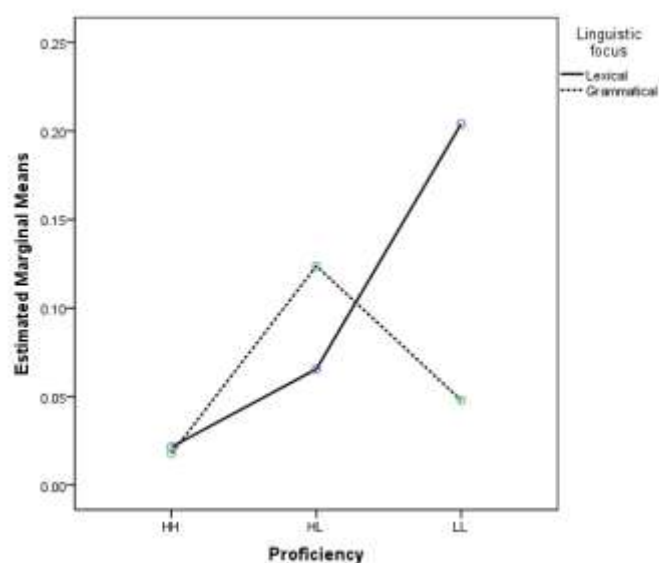
There was only a small data set for unsuccessful uptake. The RM ANOVA results showed that neither the main effect of linguistic focus ($F(1, 21) = .218, p = .646$), nor the main effect of proficiency, $F(1, 21) = .368, p = .696$, nor the interaction effect between them, $F(1,21) = .815, p = .456$ (see Figure 9.10) was significant. However, this result should be treated with caution because of the very low frequency of the unsuccessful uptake data.

FIGURE 9.9: Unsuccessful uptake by linguistic focus across proficiency groups



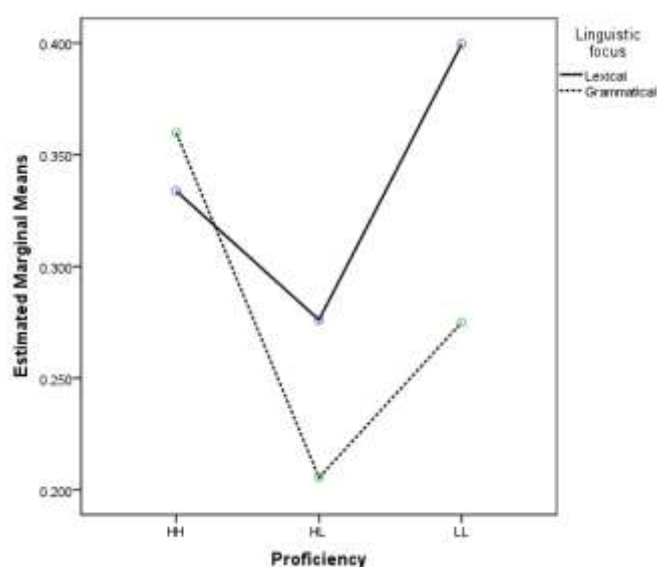
Regarding incorrect uptake, the results showed that, as in the case of successful uptake, the interaction between linguistic focus and proficiency was also significant, $F(2, 21) = 6.370$, $p = .007$, $\eta_p^2 = .38$, indicating that proficiency groups took up lexical and grammatical LREs differently (see Figure 9.10). The LL dyads had significantly more incorrect uptake when the focus of their LREs was lexical ($M = .20$, $SD = .07$) than grammatical ($M = .05$; $SD = .07$), $t(7) = 3.584$, $p = .009$, $r = .80$. In contrast, the levels of incorrect uptake lexically and grammatically was small and the same for the HH and not statistically different for the HL ($Z = -.943$, $p = .345$, $r = .33$).

FIGURE 9.10: Incorrect uptake by linguistic focus across proficiency groups



On the question of whether lexical or grammatical LREs were more likely to lead to *no uptake*, the results show that although an absence of uptake involved a higher proportion of lexical ($M = .34$, $SD = .11$) than grammatical items ($M = .28$, $SD = .24$), the difference was not statistically significant, $F(1, 21) = 1.012$, $p = .326$. The main effect of proficiency was also not statistically significant, $F(2, 21) = 1.813$, $p = .188$, nor was the interaction between linguistic focus and proficiency $F(2, 21) = .618$, $p = .548$ (see Figure 9.11). The level of no uptake lexically and grammatically for the groups were: the HH, $M = .33$ ($SD = .14$) and $M = .36$ ($SD = .33$); for the HL, $M = .28$ ($SD = .07$) and $M = .21$ ($SD = .20$) and for the LL, $M = .40$ ($SD = .09$) and $M = .28$ ($SD = .15$) respectively.

FIGURE 9.11: No uptake by linguistic focus across proficiency groups



Taken together, the linguistic focus of LREs did not affect the successful uptake for the HL and HH dyads, but it did for the LL dyads who were more successful at taking up grammatical than lexical items. Similarly, it was the LL dyads, who had a greater level of incorrect uptake when the focus of their LREs were lexical. However, when no uptake occurred, whether LREs were grammatical or lexical did not have a significant influence.

9.3 Discussion

In this discussion, I focus on the impact of tasks, proficiency and the linguistic focus of LREs on how language items targeted in LREs in rehearsal were taken up in task performance. Then I discuss the problem-solving strategies the learners employed in rehearsal in the service of task performance to further our understanding of teaching and learning through oral tasks with a rehearsal-performance approach.

9.3.1 Task effects on uptake

The findings show that the debate task led to more uptake than the problem-solving task, irrespective of dyad proficiency. There could be two main explanations for this result.

The first explanation involves the need for uptake that each task required. Although both tasks entailed a subsequent public performance, the debate task tended to pose a greater necessity for uptake than the problem-solving task. The debate task required dyad members to defend their viewpoints and negate their interlocutor's. Therefore, as dyad members prepared for their public co-performance, they needed to pay attention to matching arguments and counterarguments, thus exerting greater strain on the take-up of the ideas discussed and accordingly language items needed to convey such ideas than in the problem-solving task. In the performance of the latter task, dyad members did not need to match each other's justifications for the charity choices, thus reducing the individual need for relevant L2 words/phrases to formulate such justifications, and hence reducing uptake. Excerpt 1 (from the debate task) and Excerpt 2 (from the problem-solving task), which took place at the last minutes of the rehearsal provide support for this line of explanation.

Excerpt 1 (The debate task-HL-11a)

- S1: Cố gắng nhớ nghe. Khi tau nói ý ni, thì mi nhớ nói ý ni.
(Try to remember. When I talk about this idea, you remember to talk about this idea.)
- S2: Rồi. (OK.)
- S1: Rồi xong bơ tau nói nghèo thì lấy tiền mô ra mà làm từ thiện.
(Ok then I'll say how can poor people have money to do charity.)

- S2: Ủ, bơ tau sẽ nói có thể làm từ thiện bằng nhiều cách, rửa rửa ...
(OK, then I'll say people can do charity in many other ways, blah blah ...)
- S2: OK.
- S1: Nhớ là phải khớp ý đó nghe.
(Remember that our arguments have to be matched.)
- S1: Đương nhiên rồi. Ok.
(Of course. OK.)

Excerpt 2 (The problem-solving task-LL-11c)

- S1: Tau nói 2, mi nói 3 hi?
(I'll be in charge of option 2, and you option 3, OK?)
- S2: OK, cứ hỏi why đi rồi tau trả lời.
(Ok, [when we're up there] you'll just ask why and then I'll answer.)
- S1: Ủ. Rồi nờ hỏi bạn why và bạn trả lời.
(OK. Then you ask me why, and I'll just answer.)
- S2: Rồi. Cứ rửa thôi. Sau khi dẫn vô được rồi là cứ rửa thôi.
(OK. Just like that. Just mind the lead-in to the conversation, after that, just like that.)

Excerpt 1 shows that for the purpose of the performance, both dyad members are conscious of the need to match their points and counterpoints in the debate task, whereas in Excerpt 2, it is apparently more up to the individual dyad members to provide their justifications in their own ways in the problem-solving task, though they had earlier assisted each other to prepare linguistically for the justifications. These episodes show an advantage for uptake in the performance of the debate task.

A second explanation concerns the ways the two tasks lent themselves to the public performance. According to Duff (1986), the problem-solving task is convergent in nature, while the debate task is divergent. The need to reach a convergence on the decision-making is hypothesised to push more interaction, more language processing and more learning (see Duff, 1986; Ellis, 2003; Mackey, 2012; Pica et al., 1993). However, in the context of the current study learners collaboratively worked towards a joint public performance. Thus, it is not so much convergence or divergence *per se* that made a difference (see Chapter 8), but rather, how convergence and divergence are connected to the public performance that seems to matter. In the problem-solving

task, once the problem was solved in rehearsal, that is, when dyad members had agreed on the two charity options, the task was more likely to be perceived as being already done. In other words, the problem-solving task was probably limited in respect of an emotional expectation of a subsequent acting out or performance. In contrast, in the debate task, each dyad member had a different viewpoint that instigated back and forth argumentation, which could need to be performed to the audience. Put another way, the debate task seemed to have more performance territory as the second time learners re-engaged with the task. This was confirmed by students' perceptions of the two tasks:

- (1) *I like the task where we argued against each other (the debate task) better, as it was easier to talk about, and more fun. Importantly, it was easier and more stimulating to act out on the stage (do the performance) [laugh]. (NHS-11e-lower proficiency learner)*
- (2) *I like both tasks. However, I preferred the task where one of us is against and the other is for (the debate task) because it was interesting to argue against each other [laugh]. It was easier and more interesting to make up situations to open and close the conversation. (PDT-11f-higher proficiency learner)*
- (3) *I don't like the task that requires choosing from the five charity options (the problem-solving task). It was imposing. We tended to choose two options and just find the reasons to easily agree. The debate task? I love it! It was more fun when we perform the task in front of the class. Our classmates would be more interested in how we disagreed with each other. (BHTA-11a-higher proficiency learner)*
- (4) *Why give five options, why not let us think about our options? I preferred the arguing task. We could think more freely in the debate task, we tried our best to think of original ideas so as to deliver a nice performance later. Creative ideas are important; they show your thinking ability. (TVT-11c- higher proficiency learner)*

Three main points emerge from the comments above: i) the learners liked the debate task better because they enjoyed the content flexibility, ii) the learners, irrespective of proficiency, repeatedly mentioned the motivating power of performance and in this regard, rated the debate task over the problem-solving task and iii) the learners perceived the debate task as easier than the problem-solving task (in light of task

performance). The learners' favourable dispositions towards the debate task may explain the higher level of uptake in this task. Above all, this suggests that some tasks lend themselves to (public) performance in ways other tasks do not. This is pedagogically interesting for teachers to note in selecting tasks to implement with a rehearsal-performance approach.

Given such advantages of the debate task, the interaction effect between task and proficiency on successful uptake is interesting. The results show that while the HH and LL dyads had a greater proportion of successful uptake in the performance of the debate task than the problem-solving task, the HL dyads achieved similar rates for both tasks. This result suggests that the marked proficiency difference in these dyads maximised opportunities for collaborative problem-solving and thus overrode task effects. In other words, whether it was the problem-solving task or the debate task, the HL dyads always assisted each other the most they could, thus mitigating the effects of tasks (see 9.3.2.2).

The greater effect of the debate task on the successful uptake for the HH dyads than the LL dyads is also interesting, and seems to be in line with the presumption that closed tasks might have limited benefit for advanced learners (Lambert & Angler, 2007; Nunan, 1992; Rankin, 1995, cited in Lambert & Angler, 2007; Willis, 2004). Although both tasks used in the current study were open-ended (no single correct task outcome required), the problem-solving task was more closed or constraining in terms of content than the debate task. Regarding the creativity dimension in tasks, Rankin (1995, p.7, cited in Lambert & Angler, 2007) speculated that

A closed task ... may be so easy for participants that little if any creativity is fostered. Once a problem-solving heuristic was agreed upon by the participants, the solution and the language needed to achieve it, was highly predictable. The open task, on the other hand, allowed for considerable creativity. Opinions were shared and weighed, refined, reflected; humour sometimes played a role Closed tasks may discourage such creativity as they focus on the problem or puzzle at hand. (p.41)

Accordingly, those learners who value creativity in task performance might well be more motivated by tasks that promote it than tasks that do not.³⁵ It would be useful if future research could investigate how learner differences in creativity affects L2 learning (cf. Albert, 2011).

The fact that the debate task had more motivating power for the (public) performance than the problem-solving task could also account for why all the dyad groups had more incorrect uptake in the debate task. Incorrect uptake in this case might be called ‘successful’ uptake, as learners successfully carried the incorrect language resolutions to the performance. That learners viewed the debate task favourably may have encouraged creativity and risk-taking. It could also be that the debate task had more incorrect LRE resolutions, and therefore it had more incorrect uptake. This finding deserves attention, recollecting that the teachers (and students) in this study preferred the debate task to the problem-solving task, as they believed the former allowed greater space for the public performance (also see Chapter 5). The implication is that when teachers use debate tasks they may need to consider providing more support and feedback because learners are more likely to resolve their LREs incorrectly in rehearsal and as a result obtaining more incorrect uptake in performance (see 10.3, Chapter 10).

Taken together, the findings bring learner agency to the fore in foreign language learning (Lantolf, 2012). The learners’ task perceptions are revealing of the central role of task motivation, the extent to which a task motivates students, in language learning (Dörnyei, 2001, 2002; Dörnyei & Kormos, 2000). They also highlight the discrepancy between pedagogical and social perspectives on language teaching and learning (Allwright, 1996). From a pedagogical perspective, the greater number of LREs in rehearsal of the problem-solving task suggests that this kind of task offers more ‘occasions for learning’ (Swain & Lapkin, 1998) than the debate task. And yet it did not generate a higher level of successful uptake. From a social perspective, this could be explained with reference to the way learning opportunities were emotionally construed by participants (Allwright, 2005). The learners’ perspectives have shown

³⁵ In interviews, the learners spoke of how highly they valued creativity and tasks that allowed them to draw on their own ideas and experience (see 5.3.1.3, Chapter 5).

that it was the affective social dimension involved in the public performance that determined learning (uptake) rather the technical aspects of tasks alone. This calls for understanding what learners think and perceive about tasks, particularly in the case of public dyad performance that involves not only cognitive but also social and affective dimensions (see Batstone, 2010 for a socio-cognitive view).

9.3.2 Proficiency effects on uptake

The findings show that all proficiency groups achieved similar levels of total uptake in task performance. A focus on the learners' need to deal with language problems in order to convey wanted meanings to an audience in task performance may have cancelled out the effect of proficiency.

The focus on need was obvious given that the learners chose to initiate LREs to serve the expression of their own meanings in rehearsal. Prior research (Ellis et al., 2001a; Long, 2005; Long & Robinson, 1998; Nassaji, 2010; Williams, 2001, 2005) has shown the effectiveness of focus on form which derives from learner need. The importance of need is also in line with the involvement hypothesis on vocabulary learning (Laufer & Hulstijn, 2001) which argues that need is one of the three constructs (need-search-evaluation) that affects retention. In the present study, the level of need was high and in seeking out language solutions both from themselves and others, the level of search was, therefore, equally high. The degree of evaluation was also high, since learners had to judge which solutions fitted the meaning they wanted to express. So it could be argued that the level of involvement by the sum of the three constructs was equally high for all proficiency groups. Furthermore, the impending public performance, a central notion in this teaching and learning context appears to have added more need and motivation, and thus increased uptake of items focussed on in LREs, regardless of proficiency.

Although proficiency did not have a significant effect on the total uptake, it had distinct effects on the quality of uptake, as calculated out of the total LREs. The finding that the LL dyads, while having more LREs (Chapter 8), had a significantly lower proportion of successful uptake than the HH and the HL dyads, contrasts with previous studies which have found a positive correlation between the number of LREs and

individualised post-test scores (e.g., Storch, 2002b; Swain & Lapkin, 1998). The different results might well lie in the nature of tailor-made post-tests as a measure of learning. Post-tests of this type usually involve the testing of isolated language items with no limits on response time (Loewen, 2007, p.115), thus allowing learners to access explicit knowledge and process test items more consciously with more control than in spontaneous language use (Adams, 2007; Loewen, 2005, 2007; Nassaji, 2010). In contrast, public performance in the current study involved contextualised language use under communicative pressure, providing different evidence of learning. In this light, while the greater number of LREs produced by the LL dyads could be seen as more opportunities for learning (Swain, 1998), it may not be so. Rather it could be how learners resolved their language problems in rehearsal that affected the quality of uptake in the performance.

The fact that the LL dyads had a significantly lower level of successful uptake was not surprising, since they had a significantly lower proportion of correct resolutions in task rehearsal than the other groups (see 8.4.1, Chapter 8). What was interesting was that the HL dyads were able to achieve a level of successful uptake similar to the HH. This finding could be related to the fact the HL were able to correctly resolve their language problems equally as well as the HH (Chapter 8). A separation of the successful uptake data for the HL dyads showed that of the 62% ($M = .62$) rate of successful uptake, the lower proficiency learner ($M = .36$, $SD = .05$) had a significantly higher proportion of successful uptake than the higher proficiency peer ($M = .26$, $SD = .07$), $t(14) = -3.260$, $p = .006$, and than the average lower proficiency peer in the LL dyads ($M = .21$ (.42/2))(see Table 9.3), $t(14) = -6.472$, $p < .001$. This suggests that the learning outcome may depend on the availability of expertise. In the HL dyad, it was the higher level learner who significantly resolved more language problems than the lower proficiency peer who initiated more LREs (see Chapter 8). The higher proficiency learner responded to appeals for assistance from the lower proficiency interlocutor, or provided correction/assistance without being asked. In either case, the lower proficiency learner benefited (Examples 1 and 2) because he or she needed the language item to convey meanings to the class audience. It would seem that the LREs lower proficiency learners had when working with higher proficiency learners were

associated with more learning. It follows then that it is simply not lower or higher proficiency learners *per se*, but how they are paired together that appears to have an influence.

Example 1 (The problem-solving task-HL-11f)

Rehearsal	Performance
Low: Ngoài ra là rằng hèo? (How to say 'besides'?)	Low: [] Er In addition we can we can use this money to (.) encourage poor students to study hard and with it they can learn without
High: Besides	erm (.) worry about the ah finance of the family ah then they can learn better
Low: Besides	High: I agree with you and er I think providing (.) providing job training or education programme for disabled people is within our ability, erm and in my opinion we I think we should save this money in a bank every year
High: or in addition (.) in addition	[]

Example 2 (The problem-solving task-HL-11c)

Rehearsal	Performance
Low: Hi Tu [laugh]. Last night did you watch ah the programme?	Low: Hi Tu
High: the auction charity auction	High: Hi
Low: huh?	Low: ahm did you watch er the charity did you watch the charity auction on TV last night?
High: auction auction	High: yeah I think that's a great programme.
Low: option? Ghi rằng?(How to spell it?)	Low: after (.) after watching the programme I hope I can do charity, too. Erm if you if you have 500 million VND, what (.) how will you spend it on do (.) doing charity?
High: a-u-c-t-i-o-n. Auction a!(Auction!)	
Low: Rồi, chi? auction chi? (ok, what? auction what?)	
High: charity auction	

The results that the lower learner achieved a higher rate of successful uptake than their higher proficiency interlocutor in the HL dyads seemingly contrasts with the research by Watanabe and Swain (2007) who found novice learners in expert/novice pairs internalised less than the higher level learners (i.e., they had a lower post-test score). The different results may be, again, due to the measure of learning. These authors explained that the higher proficiency learner gave more assistance and explanations and thus processed language items in a deeper manner than their lower level peer, hence scoring higher in the post-test. In the current study, while the higher learner in the HL dyads also supplied more assistance, they did not obtain a higher level of successful uptake. Since the lower proficiency learners sought more help during rehearsal to resource their performance, it is not surprising that, provided with correct language solutions, they showed higher levels of uptake, and therefore potentially more learning.

The higher rate of successful uptake for the lower learners in the HL dyads does not necessarily mean that the higher learners in the HL did not benefit as much from interaction. They also benefited as shown by their uptake of language points attended to in LREs, though at a smaller proportion (out of the total LREs for the dyad), simply because they did not have as many language problems. But the benefits to these higher proficiency learners are likely to also include increased confidence and fluency gains through accessing and activating their language resources to resolve mostly language problems the dyads were faced with (see Chapter 8). Ohta (2001) reported that the more proficient learners when working with lower proficiency learners honed their language fluency and the 'awareness of the status of their own knowledge'. Van Lier (1996) also argued that students can learn by being the teacher for their peer.³⁶ The current data did not provide a means of measuring learning for those who provided assistance, because it was the learners who asked for or received assistance that took up the items in the public performance (see 10.6, Chapter 10).

By incorporating the proficiency variable, the current study has demonstrated that uptake of incorrect forms is not the same for every dyad. In particular, it is more

³⁶ Learners can also benefit by just being observers and listeners (e.g., Newton, 2013; Ohta, 2000, 2001).

consequential for the lower proficiency dyads (more than eight times and two times as much as the HH and HL respectively), given their more limited L2 resources. This result could be explained, once again, by the relative proportions of incorrectly solved LREs in task rehearsal by different dyads (LL>HL>HH). The result could suggest that learner-learner interaction can bring about 'wrong' learning (Swain, 1998) or 'mis-learning' (Adams, 2007), though uptake of incorrect forms was not predominant in this data (see Storch, 2002b for similar results). This raises the question of whether or not incorrect feedback that arises in peer-peer interaction should be a concern. On the one hand, uptake of incorrect peer feedback might simply indicate temporarily varied interlanguage systems (Bruton & Samuda, 1980; Gass & Varonis, 1989; Porter, 1986; Richards & Rodgers, 2001). On the other hand, as learners resort to incorrect forms across tasks under the pressure of making meanings to be conveyed publicly, the incorrect forms may enter their interlanguage without their awareness of these forms being non-target-like, leading to fossilisation. In this regard, teacher feedback and support would be useful especially when dyads with limited L2 resources work together (also see Chapter 10). Above all, collaborative rehearsal seems to shape uptake differently, depending on the learners' ability to find correct solutions to their language problems in rehearsal. The relationship between the ways learners resolved their LREs and corresponding uptake is discussed in greater detail next.

9.3.3 Relationship between LRE resolutions and uptake

The results show that whether it was the problem-solving task or the debate task, LRE resolutions had a significant positive correlation with the quality of uptake. Also, once learners found solutions to their language problems (whether correct or incorrect), they used a majority of them in the performance. However, once they left their language problems unsolved, they did not use the language items focussed on in these unsolved LREs at all. These findings are important and could be interpreted from three perspectives: socio-cultural, psycholinguistic and cognitive.

From a socio-cultural viewpoint, language knowledge is formed and developed in social interaction via language as a mediating tool (Lantolf & Thorne, 2007; Swain, 2000, 2001, 2006; Vygostky, 1978). In collaborative rehearsal, learners talked about

their language problems (LREs), and assisted each other to resolve them, thus shaping knowledge and learning (uptake). Whether their language issues were resolved correctly or incorrectly, both were associated with learning. Example 3 shows one of the dyad members maintained correctly the correct language resolutions provided by their peer.

Example 3 (The problem-solving task-HH-11b)

Rehearsal	Performance
S1: I'm erm mình nói kinh doanh have business à? (<i>I want to say 'do business', should it be 'have business'?</i>)	S1: Hi Linh. How are you doing?
S2: I do business thôi!(<i>I do business!</i>)	S2: I'm fine. And what's your job?
S1: I do business and erm I gain kiếm được ... kiếm được là chi? (<i>earn ... how to say 'earn' (money)?</i>)	S1: I do business and I earn a lot of money and I want to take uhm part in volunteer work
S2: raise (.) uhm kiếm được là chi hè (<i>how to say 'earn'</i>) (.) earn (.) earn!	S2: Ok. That's a good idea and erm what are you going to do with this money?
S1: and I earn a lot of money	

Here in rehearsal S1 wants to encode '*do business*' in English, but she is not certain about her language choice ('*have business*'). She questions her language use, which her peer (S2) responds with a correct solution '*do business*' which S1 then incorporates into her speech and moves on with her communication. However, as she moves on, she requests for assistance with another English word to express the meaning of '*earn (money)*'. S2, after having mobilised her language repertoire, resorting to L1 (Vietnamese) to mediate her find, provides a correct solution ('*earn*'), which S1 again not only uses in her immediate utterance but also later maintains in the performance. Clearly, S1 is able to perform beyond her competence right there in task rehearsal, with assistance, and independently later in the performance. This seems to offer some evidence of internalisation from interpersonal to intra-mental parameters (Lantolf, 2000, 2006; Lantolf & Thorne, 2007; Vygotsky, 1978, 1987; Wertsch, 1985) and in the process of internalisation, language or 'speech' plays a mediating cognitive role, as

learners verbalise their language issues and questions their language use (Swain, 2000, 2006; Swain & Lapkin, 1998). This also holds true in cases where students collaboratively found the resolutions to a given language issue faced with. Example 4 is one such case.

Example 4 (The debate task-LL-11a)

Rehearsal	Performance
S1: Bơ xong mi nói dễ cảm thông, bơ tau nói (<i>After that you'll say it is easy to be sympathetic, but I'll say</i>) (.) but I think, nếu (<i>if</i>) (.) if you just (.) just only you are just only sympathetic with them (.) chẳng làm chi cả (<i>it's no use</i>)... it's no (.) it's no (.) it's no use sympa ...	S1: I think it's <i>no use sympathising</i> er with er no money, I think so.
S2: sympathise	S2: I think money is important, but erm it is it is all
S1: sympathising. Mi nói trong khi đó ... but nói là money can giúp đỡ về tinh thần a tề và vật chất (<i>You'll say while ... but money can help materially and spiritually</i>)	S1: not all? S2: not all, is not all. We can visit them and talk about (.) talk about problems and share about your sorrow and joy ah and encourage them to get over their difficulty.

Again, a transition here is from rehearsal where attempts are made collaboratively from both self and the other to assist the conveyance of the intended message to performance where S1 is able to use the language item correctly, even in extended talk '*I think it's no use sympathising er with er no money, I think so*'. This rehearsal dialogue shows how the interlocutors' utterances have become an 'object' for each other to further scrutinise (Swain, 2000; Swain, Brooks, & Toccalli-Beller, 2002). In other words, the learners were able to supply each other with 'affordances' - "what is available to the person to do something with" (van Lier, 2004, p.91).

From a psycholinguistic perspective, in rehearsal, learners noticed the hole between what they wanted to say and what they were capable of saying (Swain, 1998, 2005) or the gap between their language use and the target language and made an effort to fill these holes and gaps. Since noticing is a process central to learning and acquisition

(Schmidt, 1990, 1994, 2001; Schmidt & Frota, 1986), it is then not surprising that the words used to fill these holes/gaps (whether correct or incorrect) were used in the subsequent performance. Swain (1998, 2005) argues that by producing language (oral or written), learners also form and test hypotheses about language through feedback from interlocutors. By making intended meanings, learners are ready to acquire the forms needed to make such meanings (Long, 1996, 2007, 2009; Long & Robinson, 1998). Feedback received from each other brought heightened attention to the form in use, thus leading to uptake.

From a cognitive, information-processing, perspective, the second time the learners re-engaged with the task, in public performance, the language items attended to in task rehearsal became readily available for use, because they were no longer preoccupied with what they were to say, thus having free attentional resources to attend to saying the intended content (Bygate, 1996, 2001; Ellis, 2005, 2009b; Skehan, 1998, 2009), hence facilitating uptake. It is understandable that once language items have been accessed or made available in rehearsal (for example, '*do business*' and '*earn*' in Example 3 and '*sympathizing*' in Example 4), they are likely to be taken up in the subsequent performance. This is also true for language items that resulted from incorrect LRE resolutions. The rehearsal-performance dialogues above show how effortful rehearsal was when learners were working out solutions to language problems encountered, yet in the performance they were able to use language items immediately with apparently less effort. This is also compatible with the cognitive and skills development theory (DeKeyser, 1998, 2001, 2007) which argues for the role of practice in improving and automatizing L2 use. Han (2002) suggests that "repeated practice without cognition leads only to rapidness; practice with cognition leads to improvement" (p.18). Performance following rehearsal as the second involvement with the same task was not a mere repeated practice opportunity, but an opportunity to perform the task in front of a public audience that may have both pressurised and catalysed the learners to use both correct and incorrect solutions they were able to find in rehearsal. Also regarding performance as public performance of 'high communicative pressure', under time pressure (a 15-minute rehearsal) and the demands of the tasks, the complete absence of uptake for unsolved LREs in both tasks

by all proficiency groups was because students dropped ideas that involved their failed language solutions (see Episodes 1-2).

Episode 1 (The problem-solving task-LL-11c)

- S1: healthy (.) they trong sự dẫn vật của cơn bệnh là chi hè?
(how to say 'in the torment of the disease'?)
- S2: khó quá! Thôi mà mình nói chi mấy cái phức tạp rứa. Nói chi đơn giản thôi!
(too hard! But why do we have to say such complex ideas. Let's say something else simple!)
- S1: Vấn đề là thiếu từ vựng [laugh]
(The problem is just a lack of vocabulary) [laugh]

Episode 2 (The debate task-HH-11a)

- S1: encourage them to study better xxx erm tạo điều kiện là chi? *(how to say 'create conditions'?)*
- S2: Chịu. Thôi bỏ qua đi, ý khác.
(I give up. Stop, skip that, think of other ideas.)

From any theoretical perspective, the finding that learners were able to use a majority of their correct LRE resolutions (61-74%) in the performance, was encouraging and comparable to the findings in studies that measured learning through tailor-made post-tests. For example, in the context of teacher-learner interaction, research has shown that learners scored correctly 50-60% (Loewen, 2005) and 69-73% (Nassaji, 2010) of the time in tailor-made post-tests. Findings from learner-learner interaction showed retention rates in tailor-made post-tests, of 70-80% with dictogloss tasks (Eckerth, 2008; Kim, 2008; La Pierre, 1994, cited in Swain, 1998; Swain, 1998), 59% with information-gap tasks (Adams, 2007), at least 40% with a range of activities (William, 2001), and 28-48% with computer self-accessed activities (McDonough & Sunitham, 2009).

The successful uptake rates of 61-74% of the correct language resolutions are also encouraging in view of findings from studies that gauge learning via subsequent individual presentations. Donato (1994) found that in subsequent individual

presentations one week later, a group of French ESL learners used 75% of the language items (24/32) targeted in LREs in the preceding one-hour group pre-task planning. The planning time was long (one hour) and the delayed subsequent presentation could have created conditions for learners to invest extra individual effort to prepare for it, thus enhancing retention (see Eckerth, 2008). However, in a study with similar design to Donato's, Truong and Storch (2007) found not only low incidences of LREs but also a lower uptake rate of 7.1% (1/14)-28.6% (4/14) of the language items focussed on in LREs in a previous 20-minute group planning session. These authors explained that the nature of the presentation as a context for language practice (not for evaluation), and the unguided planning caused their learners to pay little attention to form.

The performance in the current study was also not for assessment and the rehearsal was also unguided. The question arises why the results were different. Performance in pairs in the present study, which differs from the individual presentations in Truong and Storch's and Donato's studies, may account for greater occurrences of LREs and associated uptake. It was possible that dyad performance entailed greater pressure in terms of a fluent and matched conversation between the two dyad members and thus greater co-responsibility. This was confirmed by learners' verbal reports (see 9.3.5). In contrast, in individual presentations, it is more likely up to individual students to deliver the presentation in the way they want. They can use different language items to express the same content, thus reducing the chance for uptake. Furthermore, in joint performance, the interlocutors possibly had time in between turns to get ready what they were to say, thus facilitating uptake (see Michel, 2011; Tavakoli & Foster, 2008 for a discussion of how monologues and dialogues entail attentional resources). Clearly, more research into whether or not individual presentations or dyadic (public) performance creates different learning opportunities is needed.

To summarise, the types of language resolutions (correct, incorrect, unsolved) learners found in task rehearsal were closely linked to the quality of uptake in task performance and this was interpreted from multiple perspectives.

9.3.4 The linguistic focus of LREs and uptake

This section focuses on whether lexical or grammatical LREs in rehearsal were more likely to be taken up in task performance. The results showed that the LL dyads were more likely to achieve successful uptake when their LREs were grammatical rather than lexical in focus. This finding is interesting, given that grammatical accuracy is typically a lower priority when cognitive resources are stretched in face to face communication (Skehan, 1998; VanPatten, 1990). The current result does not confirm empirical findings of previous research that language features with salient communicative values were more likely to be learnt (e.g., Jeon, 2007; McDonough & Sunitham, 2009) and that learners noticed phonological and lexical items more than grammatical items (e.g., Kim & Han, 2007; Lyster, 1998; Mackey et al., 2000; Sheen, 2006). For example, McDonough and Sunitham (2009) found that their Thai EFL university learners remembered significantly more lexical than grammatical items in a post-test (48% and 28% respectively). Lexical LREs in the current study largely involved explicit requests for assistance which were often responded to with explicit answers. In other words, the answers were typically straightforward, in the form of ‘provide’ or ‘inform’ (see Ellis et al., 2001a, 2001b). Coupled with the communicative stress of a public performance, one might expect more successful uptake of lexical items to dominate. Two factors are likely to have contributed to the more successful uptake of grammatical items by the LL dyads.

First, grammatical LREs typically represented rules that the learners had already encountered previously but needed reminding or consolidating. They were more often triggered by a grammatical error for which the learners self-corrected or were corrected. In contrast, lexical LREs more likely represented new knowledge, as they were often initiated by the learners who asked for the word they needed to express the intended meaning. So if “gaps in knowledge are also a matter of degree and can relate to the extent to which students have control over a particular form” (Nassaji, 2010, p.918), grammatical LREs may be less of a gap than lexical LREs. This was so for the LL dyads who were more successful at correctly resolving grammatical than lexical LREs (see Chapter 8), thus leading to more successful uptake of grammatical than lexical items.

Second, grammatical LREs generated by the LL dyads were typically resolved with metalinguistic explanation, which has been argued to be more beneficial in sustaining noticing than LREs without it (Fortune, 2005; Storch, 2008). Episodes 3-4 show the striking contrast between how a lexical LRE and a grammatical LRE are typically resolved by the LL dyads.

Episode 3 (The problem-solving task- LL-11c)

- S1: oh that's really a big (.) big money erm I think
S2: much! much! money is not đếm được (*countable*).
S1: much (.) much money (.) much money. I think if you should spend it for yourself, just for yourself, it's so [selfish
S2: [selfish

Episode 4 (The debate task-LL-11b)

- S1: Giáo dục là chi hề?
(*How do you say 'education'?*)
S2: Education.
S1: Education.

This difference in the ways the LL dyads resolved grammatical and lexical problems may have explained why they were more successful at taking up grammatical than lexical items in the performance. In contrast, while the HH and HL dyads typically resolved their lexical LREs in similar ways to the LL dyads ('ask' and 'provide'), they often resolved their grammatical LREs by recasts (i.e., restatement of the meaning using the correct form), without explicit metalinguistic explanation, as shown in Episode 5. Research has shown that implicit feedback such as recasts are less effective than explicit feedback in the form of metalanguage explanation (e.g., Ellis, Loewen, & Erlam, 2006; Lyster & Ranta, 1997). This may explain why the HH and HL dyads did not take up more grammatical than lexical items than the LL dyads.

Episode 5 (The debate task-HH-11a)

S1: erm charity not mean

S2: DOESN'T MEAN

S1: doesn't mean you give money to help the poor

The finding is interesting and contrasts with what was found in previous research that low proficiency learners were less likely to attend to form (Kim & McDonough, 2008; Leeson, 2004; Williams, 1999, 2001) and explain their language resolutions metalinguistically (e.g., Fortune & Thorp, 2001). This seems to show that it is not necessarily that low proficiency learners are less capable of making form-meaning connections. But rather the tasks used and task conditions (e.g., rehearsal-performance) may play a catalysing role. Given a rehearsal-performance condition, learners of differing proficiency levels or L2 resources addressed their own language problems, whether grammatical or lexical. Thus they addressed their own needs and shaped subsequent learning.

To further understand uptake in task performance, in the next section, I discuss learner agency, a hidden dimension behind the 'private' rehearsals that show the problem-solving strategies learners of differing levels of resources (proficiency) chose to employ to deal with the pressure of the performance.

9.3.5 Problem-solving strategies: Proficiency and performance pressure

Underlying the evidence of uptake were the problem-solving strategies that the learners undertook in task rehearsal to prepare for task performance. These strategies were shown through their rehearsal talk and reported perceptions.

The rehearsal discourse showed that the lower proficiency learners used more remembering and rehearsing strategies than the higher proficiency counterparts, who rather took a casual approach and some even preferred improvising. For example, explicit mentions of remembering, as in Excerpts 3-5, were more typical of low proficiency learners.

Excerpt 3 (The debate task-LL-11e)

S1: Cố gắng nhớ, mi làm đừng có quên.

(Try to remember, don't forget.)

S2: OK. Ê lại, có nhiều từ tau không nhớ.

(OK. Hey again, there are many words I can't remember.)

Excerpt 4 (The problem-solving task-HL-11c)

High: Rứa là vô rồi đó!

(Good, we can lead into the conversation already!)

Low: Tau sợ tau quên.

(I'm afraid I will forget.)

High: Quên đồ rứa nữa à?[laugh]

(That involves forgetting and remembering for you?) [laugh]

Low: Ừ, dễ quên lắm mi ơi. Tau phải nhớ mấy từ ni.

(Yeah, very easy to forget. I have to remember these words.)

Excerpt 5 (The problem-solving task-LL-11c)

S1: Cho bạn xin tờ giấy, bạn cần phải nhớ cái ni.

(Give me a piece of paper, because I need to remember this.)

S2: Tau nói cái ni xong rồi bớ tau nói do you agree with me.

(After I have said this, I'll say 'do you agree with me'.)

These episodes show the lower proficiency learners made a deliberate effort to remember the ideas, words or phrases they had discussed in preparation for the performance, thus promoting uptake. Storch and Wigglesworth (2010) also showed that learners improved their post-test scores because they memorised the feedback information.

Local rehearsing at phrase/sentence levels was also typical of the lower level learners.

Excerpts 6-7 are illustrative of these strategies:

Excerpt 6 (The problem-solving task-HL-11e)

Low: Oh, oh, yes (.) Er I think er I think it's very perfect (.) Er er

High: Perfect rồi còn very chi nữa mi! (*Perfect already, no very!*)

Low: I think it's perfect ... It's perfect ... It's perfect ...[Repeating to herself]

Excerpt 7 (The problem-solving task-LL-11c)

S1: (...) Hỏi why đi để bạn trả lời. (*Ask why and I'll answer.*)

S2: Hey?

S1: Hỏi why đi để bạn trả lời. (*Ask why and I'll answer.*)

S2: Why?

S1: Eh er in central Vietnam, there is (.) there are many floods ah special Hue, (...), bớ nớ trả lời (*then you continue*).

In Excerpt 6, the low proficiency peer keeps repeating the phrase '*it's perfect*' that she obtains as feedback from her interlocutor. From a socio-cultural perspective, this is a kind of private speech that mediates learning (Appel & Lantolf, 1994; Lantolf, 2000, 2012; Lantolf & Thorne, 2007). Previous research has shown that private speech can function as rehearsal (De Guerrero, 1994, 1999) or self-regulating (DiCamilla & Antón, 1997) and that private speech and internalisation are associated (see Lantolf, 2000). In Excerpt 7, the two learners rehearsed locally by practising asking and answering questions. These overt rehearsing strategies may have enhanced the learners' remembering of the ideas and associated linguistic resources to express their intended meanings when they performed the task.

In contrast, explicit mentions of improvising were more typical of higher proficiency learners, as displayed in Excerpts 8-9.

Excerpt 8 (The debate task-HH-11a)

S2: (...) Sống trong đời sống cần có một tấm lòng (*in this life what we need is a heart*)
[singing a popular Vietnamese song]

S1: Nói! (*Speak!*)

S2: Nói cho rồi, xí tau bịa sau được.
(*just finish off, I'll improvise later.*)

Excerpt 9 (The debate task-HL-11b)

- High: Hey, rứa được đủ rồi. Tau tự phịa sau. Vấn đề của tau là tau chuẩn bị chi đi nữa tau cũng rất có thể trình bày khác đi khi tau lên đó.
(Hey, that's enough. I'll improvise later. My problem is whatever I prepare, I might do differently when I am up there.)
- Low: Tau phải chuẩn bị, phải tập thôi mi nà. Mi phải giúp tau
(Me, I need to prepare, I need to practise. You must help me)

In Excerpt 8, the last utterance, *'just finish off, I'll improvise later'* that one of the higher proficiency learners verbalised, shows casualness in her approach to preparing for the performance. Excerpt 9 further shows that while the higher proficiency peer preferred improvising due to a perceived lack of transfer from planning and preparation (cf. Ortega, 2005), the less proficient learner was worried and voiced she needed to prepare and practise for the performance.

The interview data further confirmed that the lower proficiency learners made more deliberate effort in task rehearsal. Some of the lower proficiency learners stated:

- (5) *The problem for me is I lacked English vocabulary to express my own ideas. I often asked my friend. If a word is new to me, I try to memorise it; otherwise when I am up there I will forget, and the conversation might stop, and I don't want that to happen. My friend is better than me, even if he forgets a word, he can use other alternatives, and keeps on talking. I am not as good as him, so I have to remember them and repeat them or write them down in a paper just to help me remember, because no notes are allowed when we are up there. (LTKT-11d)*
- (6) *You know, within some minutes, you cannot remember everything, but you can once you make efforts to remember them. Also, I would feel bad if I forget what both of us have agreed on and practised. It is not my individual performance, but the performance for the two of us. You cannot let your partner down when he or she tried hard to help you with the words you don't know. Furthermore, I also want to be fluent on my part. (TVKT-11c)*

These comments reveal that the lower proficiency learners were concerned about their limited proficiency and automaticity, and thus decided to memorise and rehearse language items that they felt difficult to carry to the performance (cf. Ortega, 2005). Furthermore, pair performance oriented them to take more responsibility for co-delivery. It seems that performance posed greater pressure for the lower proficiency learners, and thus they took remembering strategies to serve it. Their effort to memorise their language resolutions enabled the LL and HL dyads to retain the information from LREs as much as the HH dyads (evidenced by no significant difference in the total uptake between proficiency groups). The higher proficiency learners took a somewhat different stance:

(7) *This is a joint performance by both of us, you know. So if any of us finds certain parts of our preparation or certain words or phrases hard to remember, we feel it's our responsibility to find ways to memorise them. For me it is ok. I can even improvise ideas, no problem, so I don't usually expend efforts in remembering at all, or I don't usually even take notes. Sometimes I want to express certain ideas differently from what I have prepared, and it is fun.* (PQTN-11b)

Once again the higher proficiency learners expressed a more casual approach to their task rehearsal than the lower level learners did. They were aware of their adequate automaticity, obviating the need to memorise and instead expressed a preference for spontaneity. Performance did not seem to pose as much pressure for them as it did for the lower level learners. These different strategies reflect the learners' meta-cognitive awareness and "the special thoughts and behaviours, steps or techniques that students employ often *consciously* to improve their progress in internalizing, storing, retrieving and using the L2" (O'Malley & Chamot, 1990, p.1, italics added). The public performance was problem-posing for the low proficiency learners who had lower resources. As a result, they used more rehearsing and memorizing strategies. However, the high proficiency learners who had higher resources were more confident in the performance, and used fewer of these strategies. In other words, uptake in performance for the lower-proficiency learners was more hard-won than for the more proficient learners, which has implications for pedagogy (see 10.3, Chapter 10).

Viewing uptake in light of the differing strategies that higher and lower proficiency learners used provides fuller insight into the learning process that arises from task work in a rehearsal-(public) performance condition. These differing strategies represent the learners' problem-solving approaches, and their goal-directed learning. In this way, uptake or learning in this rehearsal-performance condition, seems to be in accordance with Lantolf and Thorne's (2007, p.218) point that "what is called incidental learning is not really incidental. It is intentional, goal-directed, meaningful activity. From the SCT (socio-cultural theory) perspective, there are no passive learners and there is no incidental learning".

On the one hand, LREs were incidental in the sense that no forms were determined in advance, and they occurred in the context of doing communicative tasks (Long, 1991; Ellis et al., 2001a, 2001b; Ellis et al., 2002). On the other hand, they were not incidental in Lantolf and Thorne's (2007) sense. Students initiated LREs in the service of the higher goal of performing the task in front of the class. While LREs involve cognitive processes such as noticing, hypothesis-forming and testing (Swain, 1995, 2005), these cognitive processes cannot be seen inseparably from the social dimension that involves learners' orientation, agency, and task motivation (Batstone, 2010) under a rehearsal-performance condition. The results thus illuminate the importance of examining tasks in conjunction with task conditions and task participants. Need and problematicity do not lie only in the nature of the task *per se* or the nature of proficiency *per se*, or task conditions (e.g., rehearsal-performance) *per se*, but in an interdependence between them. The differing problem-solving strategies that the lower and higher proficiency learners used were revealing of learner agency or orientation (Batstone, 2002, 2005) which involved the learners' self-regulation of behaviour (Lantolf & Thorne, 2007) and their response to the social, interactional, interpersonal dimensions of the dyad (public) performance. Explanation thus requires complementary accounts from both social and cognitive dimensions (also see Batstone, 2010; Swain & Deters, 2007). The current results add new understandings to task-based learning, and support the need for cross-perspective understanding of task-based interaction.

9.4 Summary

The findings have shown that learners frequently engaged in addressing language problems (LREs) encountered during rehearsal to prepare for their public co-performance. They also correctly used in task performance encouraging proportions of language items which had been the focus of LREs in task rehearsal. Evidence also shows that tasks and proficiency affected not only the occurrences of LREs, and how they were resolved in task rehearsal, but also how they were taken up in task performance. Importantly, it was how language problems were resolved rather than how often they occurred in rehearsal that was closely associated with uptake in performance. The linguistic focus of LREs also interacted with proficiency to affect uptake. Above all, the motivational power of tasks in catering for the performance, as perceived by the learners was crucial in understanding the effects of tasks on uptake. The learners' orientation towards the performance and their employment of differing problem-solving strategies in rehearsal to serve the performance added more insights into EFL learning through oral tasks in a rehearsal-performance condition. As a central notion to these EFL high school classrooms, performance involved not only cognitive but also social and affective dimensions. In addition to being a catalyst and a social classroom event, (public) performance can also be used as a tool to measure uptake.

The next chapter brings together the findings of the two phases of the research and addresses implications for pedagogy, methodology, and theory as well as limitations of the study and future research areas.

10.1 Introduction

In this chapter, I present a summary of the main findings of the research, and the pedagogical, methodological and theoretical implications of these findings. I also discuss the limitations and suggestions for further research. I conclude with my contextual and personal reflection.

10.2 Summary of findings

The main findings are summarised for each of the two research phases that make up the thesis.

10.2.1 Phase 1

Phase 1 investigated (1) how EFL teachers in a Vietnamese high school implemented (oral) textbook tasks from a series of new English task-based textbooks; and (2) how students engaged in the classroom tasks. Forty five classroom observations were made of nine teachers in nine classes, each carrying out a sequence of five textbook lessons. Teachers and students were also interviewed and student task talk collected. The results for (1) are summarised in Table 10.1.

The findings showed that the teachers diverged considerably from the textbook tasks. They frequently adapted and replaced textbook tasks in preference for open-ended tasks (open, input-independent, divergent) and tasks that were real to students and situated in their immediate interests and concerns. These task choices were guided by the consistent belief among the teachers that, in order to optimise students' engagement with tasks, teachers need to connect the students with the tasks socio-affectively.

Tasks as implemented by all the nine teachers were represented in four stages: pre-task, rehearsal, performance, and post-task. Performance was perceived by both teachers and students as a driving force for target language use in the classroom. The notion of performance was central to these EFL classrooms, not only as a way of

pushing the use of English rather than Vietnamese L1, but also a social classroom event. In order to prepare for that performance, the nine teachers observed varied practices at the pre-task stage. While many teachers thought providing a model of task performance or useful language was facilitative, many students perceived it as constraining. These students also said they highly valued freedom to do tasks in their own ways.

TABLE 10.1: Teacher use and implementation of textbook tasks

	Action	Teacher thinking (why)
Task design features	(1) mainly adapted and replaced textbook tasks	To engage students socio-affectively
	(2) preferred open, divergent, input-independent tasks	To motivate the public performance of the tasks
	(3) used tasks that are real to students, their personal life and immediate interests/concerns	
Implementation procedures	<u>Pre-task</u>	
	-Varied practice: from minimal task introduction to provision of useful language, models of task performance and pre-communicative activities	To empower or to facilitate task performance
	<u>During-task</u>	Performance was seen to motivate and act as a catalyst for target language use.
	<i>Rehearsal</i> followed by (public) <i>performance</i> in dyad/groups	
	<u>Post-task</u>	
	Checked listening comprehension of the performance	To engage the whole class
	Gave feedback	To focus explicitly on form
	Discussed language issues	

The learners also said they highly valued learning opportunities through tasks in a rehearsal-performance condition. Preliminary analysis of their task talk showed that they engaged in preparing for the task to go public. They attended to language use and

helped each other to express their intended meanings in task rehearsal. Although they used Vietnamese substantially (more than 50%) in task rehearsal, they used it mainly to mediate solving language problems and to resource the subsequent performance. In the performance, they used only English. Uptake in the performance, of language points focussed on in LREs in task rehearsal, was successful at 49.4% (39/79) of the LRE cases. However, the data set contained only a limited number of matched rehearsals and performances (15), by different groups of students carrying out different tasks. This motivated Phase 2 of the research.

10.2.2 Phase 2

Phase 2 examined in more detail the extent to which learning opportunities operationalised as LREs that arose in dyadic task rehearsal were taken up in the dyadic performance of the same task. It was a focussed mixed design study that investigated the effects of task type and learner proficiency on uptake in task performance of LRE-specific items. Data were collected from 24 dyads, eight for each proficiency group (HH, HL, LL) carrying out two tasks: one problem-solving task (convergent) and one debate task (divergent). The learners were first given 15 minutes to do the rehearsal. Later they were called on to perform the task in front of the class.

The findings show that the Vietnamese EFL high school learners discussed more LREs in the rehearsal for the performance of the problem-solving task than the debate task, irrespective of proficiency. They also correctly resolved a higher proportion of their LREs in the rehearsal of the problem-solving task, though the difference was not significant. However, they achieved more uptake in the performance of the debate task, regardless of proficiency. Interviews with the learners revealed that they enjoyed the debate task more than the problem-solving task because they said the debate was less constraining in terms of task content and thus easier. They also said it had more motivating power for the subsequent public performance. That is, the debate task gave students more territory to 'act out on the stage'.

As regards the effects of proficiency, the study found that lower proficiency dyads produced more LREs in rehearsal than higher proficiency dyads, but resolved them less successfully, with a higher proportion of LREs resolved with a non-targetlike outcome.

Although proficiency did not influence the total uptake, the more proficient dyads were more likely to achieve successful uptake and less likely to produce incorrect uptake (i.e., uptake of incorrectly solved language items). It was how LREs were resolved in task rehearsal that correlated positively with the success of uptake in task performance.

Proficiency also influenced the problem-solving strategies that the learners adopted to prepare for the public performance. The lower proficiency learners typically made a more deliberate effort to remember LRE resolutions, by means of rehearsing and memorizing targeted words and fragments. In contrast, the higher proficiency dyads rarely used these strategies. Instead, some explicitly said they preferred improvisation when they performed the task.

In terms of the linguistic focus of LREs, the results showed that learners engaged in more lexical than grammatical LREs, irrespective of tasks and dyad proficiency. While the HH and HL dyads correctly resolved both grammatical and lexical LREs equally well, the LL dyads were better at correctly resolving grammatical than lexical LREs and their uptake was more successful for grammatical than lexical items. The LL dyads also had a significantly higher proportion of incorrect uptake when the focus of their LREs was lexical. In contrast, for the HH and HL dyads, the linguistic focus of LREs did not influence the level of uptake, whether successful or incorrect.

10.3 Pedagogical implications

The findings from the research have implications for pedagogy in two main areas: task design and task implementation.

10.3.1 Task design

The findings of the current research suggest four main implications for task design. First, the results suggest that in this high school context, tasks can best achieve learning outcomes if they engage students affectively and socially, by allowing them freedom to express their own meanings. To this end, open divergent input-

independent tasks are recommended, especially for motivated learners who value creativity in language use and task performance.

Second, the research findings suggest that tasks can engage students by addressing their *immediate* interests and concerns. In the English textbooks for Vietnamese high school students, the textbook authors (Hoang et al., 2007, p.7, translated into English) contend that “the topics, ... and communicative activities in the textbooks ... are selected and designed based on the principle that ... they suit students’ psychology, age, knowledge, needs and interests.” Yet we have seen the teachers adapting and replacing textbook tasks in preference for their own tasks that they believed would better engage students through topics of even more immediate relevance to the learners. By implication, curriculum designers or textbook writers in the context of a top-down centralised national syllabus as in Vietnam may need to take into account students’ and teachers’ voices in task selection and design. Task advocates argue that tasks are motivating when they are selected based on needs analysis of *real world target* tasks, tasks that learners undertake in the real world using L2 (Long, 2007; Long & Crookes, 1992; Van den Branden, 2012). However, such needs analysis may bear less relevance in an EFL context such as Vietnam. In such a context, tasks may be more motivating when they are also real to the learners themselves. This may be best achieved by the teachers who identify context-specific task design (and implementation) options that will motivate their students to participate actively in tasks. Therefore, teacher education and training that caters for teacher autonomy and thinking would be useful (also see 10.3.2.4).

Third, the results showed that both problem-solving (convergent) and debate (divergent) tasks are valuable in different ways. In a rehearsal-performance condition, the problem-solving task pushes learners to question and reflect on their language use to express the content that they agree on, thus developing and expanding their lexicogrammar in these ‘fields’ (Bygate & Samuda, 2009). This is seen in the greater number of LREs generated in the rehearsal of this task than the debate task. For this reason, teachers can add problem-solving tasks to their teaching repertoire, even though, according to the teachers and students, these tasks were more constraining. In contrast, the debate task did not induce as many LREs as the problem-solving task, but

being more open and requiring arguing back and forth, it had, according to students, more motivating power for the public performance. The debate task also pushed learners to find language solutions rather than leave them unsolved. It also led to more uptake in task performance. This suggests the debate task may be more effective for developing fluency and automaticity.

Candlin (2001) posed the question that “If the tasks are to develop learners’ meaning-making capacities ..., what is the relationship between this objective and the necessarily concurrent development of learners’ processing capacity?” (p.239). The answer is that since meaning-making stimulates searches for L2 resources and deliberations over forms to express intended meanings, what is learnt, as Bygate (1999b, p.34) claims, is “both language content and processing capacities”. In this respect, as the findings of this research have shown, both problem-solving and debate tasks contribute to learning. Both are open-ended tasks (not closed tasks that require a single task outcome), and this suggests the need for a re-evaluation of the pedagogical potential of open-ended tasks in task-based instruction.

Fourth, the findings show that teachers preferred open-ended tasks that were divergent and input-independent because these tasks had more potential to motivate the subsequent public performance. Although public performance may arguably belong to implementation, not task design features, design decisions should be informed by the (public) performance potential of a task. Open tasks with interesting topics rather than closed tasks with tight structure and fixed information provide more opportunities for unique performances. Because convergent and divergent tasks can both be open tasks, both can realise this potential.

10.3.2 Task implementation

The findings show that the teachers had all adopted a rehearsal-performance approach to implementing the textbook tasks. Furthermore, in task rehearsal, students frequently engaged in discussing and resolving language problems they encountered, and subsequently used many of these items accurately in the public performance. These findings show how effective the rehearsal-performance approach can be to task implementation for driving language development forward. Such an approach

provided ample opportunities for learners to attend to language form as they pursued meaningful task goals. In an EFL context like Vietnam and many other EFL settings where students share a native language, and have limited opportunities to communicate in the target language outside the classroom, a rehearsal-performance approach may be particularly useful to push learning through oral tasks. Although students used their native language substantially during rehearsal, they used it as a mediating tool to resolve language issues and to resource their upcoming public performance which was always entirely in English. This, in the rehearsal-performance approach to task implementation, shows L1 use as a valuable resource rather than as a major challenge to EFL task implementation as reported in some research (e.g., Bock, 2000; Carless, 2002, 2003, 2007, 2008; Eguchi & Eguchi, 2006; Li, 1998; see Butler, 2011 for a recent review). These results provide further encouragement for teachers to rethink the commonly-held doubt that whether students learn anything when doing pair/groupwork due to their homogeneous mother tongue as reported, for example in McDonough (2004).

However, for the rehearsal-performance approach to be effective, the results suggest a number of additional recommendations which are outlined below.

10.3.2.1 Forms of performance

The findings show that the teachers preferred dyadic or group public performance to public report or individual presentation. According to the teachers (and students), dyadic performance engaged students better by holding both dyad members responsible for the co-performance, thus pushing them to cooperate and prepare for it (see Chapters 4 and 5). The classroom observations also show that students created animated performances by making up opening and closing scenes to act out on the class 'stage', thus involving the whole class as the audience. The whole class was also engaged in listening to their friends' performances. This suggests dyadic or group public performance may allow for more 'acting' territory for the performance than public report or individual presentation. Further empirical evidence for the advantages of dyadic performance over public report is needed (see 10.6).

10.3.2.2 Student pairing

The findings show that task rehearsal allowed learners at different proficiency levels to adopt a focus on form and an approach to rehearsal appropriate to their particular proficiency-driven needs. The nature of LREs produced by learners at different proficiency levels tells the teacher about the particular demands the task makes of stronger and weaker learners. This then provides the basis for follow-up post-task activities and also informs subsequent task design and implementation. For example, when two low proficiency learners worked together, they met more language problems in task rehearsal. They were also more likely to problem-solve incorrectly than the other dyads (on average 2.5 times in every 10 LREs compared to 0.4 out of 10 LREs for the HH dyads and 1.3 times for the HL dyads) and they were more likely to produce incorrect uptake (1.7 times out of 10 LREs, compared to 0.2 and 0.7 times out of every 10 LREs for the HH and HL respectively). This raises the question for teachers of how best to arrange dyads. On the one hand, the LL dyads worked effectively and appropriately to navigate the task and supported each other well. On the other hand, they had a higher proportion of incorrect LRE resolutions. Being aware of this, teachers may need to provide support for the learners. This may take the shape of being available during rehearsal or providing follow-up feedback after the performance. In this way, tasks in a rehearsal-performance condition may provide what East (2012) contends as, ‘assessment for learning’ or ‘feedback and feedforward’.

The finding that when paired together, higher proficiency dyads were most successful in resolving their language problems and achieving successful uptake seems to suggest this pairing is most beneficial. However, the findings also show that when two higher proficiency learners worked together, they were more casual in their rehearsal, because the public performance did not push them as it did low proficiency learners. This points to the question of how to motivate these proficient learners to further stretch their ‘upper limits’. Increasing task complexity along the resource-directing dimension (Robinson, 2001, 2007) may accomplish this. For example, for the problem-solving task (see Appendix 8), the HH dyads could be asked to choose three charity options rather than two. Supplementary fluency-development activities can be included such as asking high proficiency learners to pick up a topic and talk about it

within gradually reduced amounts of time (Nation & Newton, 2009). Furthermore, switching task interlocutors for the public performance (e.g., a dyad member from a HH dyad performs the task with a dyad member from another HH dyad) may also push proficient learners to perform at a higher level.

That the HL dyads were able to resolve their language problems as well as the HH dyads and achieved a successful uptake level similar to the HH dyads suggests that pairing learners of a marked discrepancy in proficiency is also useful. The lower proficiency learner in the HL dyads achieved a higher level of successful uptake than the higher proficiency peer. This was simply because the latter did not have as many language problems (see 9.3.2). The higher proficiency learner typically undertook the expert role, assisting the less proficient peer with language issues as they were preparing for their co-performance. Research (e.g., Van Lier, 1996) has shown learning benefits when learners 'teach' their peers (see 9.3.2). The implication is that for mixed proficiency dyads to work efficiently, teachers may need to make sure the more proficient student understands the value of assisting the less able peer, so that she/he maintains her/his collaboration throughout rehearsal, because research (e.g., Kowal & Swain, 1994, 1997; Yule & Macdonald, 1990) has shown that if the higher proficiency learner dominates or ignores the less proficient peer, this pairing is not conducive to learning. This was not the case in the current study.

10.3.2.3 Pre-task and post-task work

Phase 1 of the research found that many teachers provided pre-task modelling and some pre-task language, because they believed these would be facilitative of student task performance. However, a majority of the students interviewed said they found these steps constraining and inhibiting. Students instead said they wanted to create their own unique performances that could impress their audience. It follows that if students value creativity in language use and task performance, (if provided at all), modelling and language input need to be provided in a way that gives room for this creative impulse. Alternatively, organised feedback or cycles of activities at the post-task stage that respond to students' earlier exploratory language use would be

recommended. For teachers to make informed decisions on this issue, it may require dialogue with students to find out what pre-task actions motivate or demotivate them.

With respect to post-task work, although there is a general consensus that teacher feedback on student language use at the post-task stage is valuable (see Willis, 1996), post-task work in the context of a rehearsal-performance approach may need to incorporate activities that motivate students and sustain their attention at the performance stage. For example, the present findings show that, to motivate students to listen to their friends' performances, the teachers checked students' listening comprehension and elicited peer comments and feedback on not only language use but other aspects of performance. The teachers also gave credit not only for the performance delivered but also for good peer comments and feedback (see Chapter 5). However, because students can be very competitive, they can give 'harsh' comments on their peers' performance, which could lead to demotivation. In such situations, teachers need to provide clear guidelines for giving supportive feedback and model such feedback themselves.

10.3.2.4 Teacher thinking

The research findings show the powerful influence of teacher thinking on task pedagogy. The teachers made decisions autonomously to adapt and replace prescribed textbook tasks based on rationales that they could clearly articulate (Chapter 5). That they did this in a Vietnamese high school, a context where the education system is centralised through a national top-down curriculum and embodied in a series of prescribed textbooks not only for English but also for other subjects, is a tribute to their professional confidence and judgement. This has implications for language policy makers, teacher educators, and textbook writers in Vietnam. Teachers need to be encouraged to use the textbook as a resource to suit the learners that they are teaching, but not to treat the textbook as a straitjacket. For textbook designers, this suggests the need for design that allows teachers more opportunities to adapt and deviate from the textbook. If we "view teaching as an activity that creates learning opportunities and learning as an activity that utilises those opportunities" (Kumaravadivelu, 1994, p.33), then the role for the teacher is directly vital in both

making available learning opportunities and assisting utilisation of them. By being given greater freedom to use mandated textbooks as a resource, teachers may be enabled to orchestrate their teaching that accommodates the '*particularity*', '*practicality*' and '*possibility*' of the teaching context they find themselves in (Kumaravadivelu, 2001, 2006a).

10.3.2.5 Language development

The present findings suggest the need to cater for student language development, not as one-off event but as an on-going emergent process (Spada & Lightbown, 2008, p.182). The findings show that learners, especially low proficiency learners, deliberately employed rehearsing and memorizing strategies to remember language items and used them correctly in the performance. This evidence of uptake in task performance is clearly only a step in the learning process that culminates in the integration and automatisisation of the targeted items. To achieve this, however, extended opportunities for language use and analysis will be needed (see DeKeyser, 2001, 2007). Additional to public performance, teachers may consider asking students to reflect on their performance and complete a learning log in which learners record items they 'discovered' in the task. Teachers may also consider using delayed task repetition (Bygate, 1996, 2001) and with a different partner (Bitchener, 2004) to afford opportunities for students to re-engage with previously accessed and activated language resources in a different way. Alternatively, asking students to re-do the task in writing may also be useful. For example, based on the debate and problem-solving tasks students have carried out orally, teachers could ask students to write individual essays (see Brooks & Swain, 2009; Storch, 2002a, 2002b). This could be done as homework due to the usual classroom time constraint. In this way, students may bring together the learning of language items processed in the earlier oral engagements with the task (rehearsal and performance), and thus process and reflect on their language use more deeply (Swain & Lapkin, 1998, 2001). These "cycles of analysis and synthesis" enable re-structuring and reorganisation of the language system, and the grounds for exemplar-based learning to become rule-based learning and vice versa (Skehan, 1998).

10.4 Methodological implications

This research has several methodological implications with regards to research design and data analysis.

10.4.1 Research design

A strength of the present research lies in its design to focus on tasks in classrooms, on both teaching and learning, and from both teacher and learner perspectives. This ‘researched pedagogy’ approach (Bygate et al., 2001) has been able to document not only task as work-plan but also task as process (Breen, 2009; Coughlan & Duff, 1994). As Samuda and Bygate (2008) argue, “without attention to both the plan and the process, the pedagogic principle is lost” (p.66).

In addition, the research (Phase 2) has a number of methodological design features which are unique to the area of LRE-based research. These include: i) keeping task topics constant across proficiency and task groups; ii) keeping rehearsal time constant across proficiency and task groups; iii) allocating students to proficiency groupings based on both objective (English scores) and subjective (teacher ratings) measures; and iv) using a mixed design with task type as a within-subject variable and proficiency as a between-subject variable, thus enhancing the validity of the claims made on task and proficiency effects.

10.4.2 Data analysis

The research has also informed task research through (1) extending the analysis of task design features, (2) expanding the ways that learning through oral tasks is measured, and (3) analysing data from multiple perspectives.

Regarding task design features, the research has highlighted the importance of two task features not yet categorised in the task literature. The first feature is degree of input-dependence (input-dependent vs. input-independent) which refers to whether students are required to use the input provided to complete the task or not (see Chapters 4 and 5). The results show that degree of input-dependence impacted on the meanings to be made and thus the *productive* LREs students discussed and the ways

students perceived a given task as constraining or not, which in turn affected their self-reported engagement in tasks, and subsequently uptake.

Similarly, the findings suggest that task authenticity needs to be viewed beyond categories of situational or interactional authenticity (Ellis, 2003). Task immediacy and personalisation are also important dimensions of task authenticity. Notably, the findings show that task authenticity is *situated*, and bears on context-specific meanings, as explained and justified by the teachers who perceived task authenticity as how 'real' a given task is to the learners in that context (see Chapters 4 and 5). This highlights the need to gain emic insights by interviewing teachers and learners to understand the situated-ness of task authenticity.

Another methodological implication involves the process-product approach (also see Donato, 1994; Storch, 2002a, 2002b; Truong & Storch, 2007) that the research employed to first identify LREs in task rehearsal and then trace the use of language items targeted in LREs in the subsequent public performance. This process-product approach has a number of advantages. First, it offers an alternative way of measuring learning that does not rely on a pre-test-post-test design that is subject to the unpredictability of LREs. Second, measuring learning through uptake in task performance offers a more spontaneous productive measure of learning compared to discrete tailor-made post-tests (see Chapters 3 and 7). Finally, by using a process-product approach to capture LREs in rehearsal and the uptake episodes in task performance, the research has been able to show how learners were pushed to expand their linguistic resources. This has overcome the one-for-all measure of the product via complexity, accuracy, and fluency (CAF) as predominantly used in planning studies (Batstone, 2005). According to Batstone, the CAF framework fails to differentiate learners who have already reached a certain level of CAF, and those who have been pushed "rigorously and deliberately" through being given an opportunity to do pre-task planning (p.280). In the present study, by capturing the process (LREs), we can see that being given an opportunity to do a rehearsal, a type of planning (Ellis, 2005) for their public performance, learners were pushed to stretch their language resources.

Finally, the data were analysed both quantitatively and qualitatively, both statistically and micro-genetically, and from both etic and emic perspectives and under the lens of both teachers and students. These multiple approaches to data analysis provide an enriched view on teaching and learning through oral tasks and add explanatory power to the findings.

10.5 Theoretical implications

Two main theoretical implications that can be drawn from this research involve the theoretical construct of tasks and the role of ‘pushed output’ from multiple theoretical perspectives.

The present findings have shown that governed by the overriding concern with learner engagement, the teachers looked to both task design features and task implementation conditions (e.g., rehearsal-performance) to engage students and push them to use the target language. This challenges earlier claims that because the construct of task by definition involves both goal and means, task design and methodology in TBLT are one and the same (e.g., Nunan, 1989). In view of the findings in this research, tasks in themselves may not be enough to push learning through oral tasks (also see Bygate, 2005; Skehan, 1998; Skehan et al., 2012; Willis, 1996). An implementation or ‘methodological’ option such as rehearsal for performance can have catalysing power. Furthermore, the teachers, by using different pre-task approaches from minimal task introduction to provision of some language input and models of task performance could facilitate or constrain task performance, as *perceived* by the learners who carried out the task. In these cases, tasks were purely content and teachers realised them in a range of ways. This echoes the view that tasks should be seen as “curricular content rather than a methodological construct” (Kumaravadivelu, 2006b, p.65) or as a pedagogical construct, a pedagogical tool for different purposes (Samuda & Bygate, 2008). Seeing tasks in this way may allow a more flexible and less dogmatic approach to TBLT, and thus more teacher ‘appreciation’ of tasks (Ellis, 2003, 2009a; Samuda & Bygate, 2008).

The present findings also reinforce the role of ‘pushed output’ (Swain, 1993, 2005). Swain (1993) argues that learners “*need to be pushed to make use of their resources;*

they *need to* have their linguistic abilities stretched to their fullest; they *need to* reflect on their output ...”(p.160, italics added). All these needs were addressed through particular design and implementation decisions made by the teachers in this study. Tasks can ‘direct’ attention to form by creating opportunities for “meaning (concept)-form (language) mapping” (Robinson, 2011b, p.14, also see Bygate & Samuda, 2009), when carried out in a rehearsal-performance condition. The learners made explicit and conscious effort to verbalise and resolve language problems in their meaning-making in task rehearsal to prepare for the public performance.

Above all, the findings supported the view that L2 learning is a more *conscious* process than has been claimed (Pica, 2005). By supporting this view, the results have strengthened the need to see how L2 learning is manifested as a conscious process from multiple perspectives. The findings have shown: (1) the importance of conceptual demands of tasks and task conditions (rehearsal-performance) (cognitive); (2) the value of noticing the ‘holes’ and ‘gaps’ and feedback (psycholinguistic); and (3) the importance of collaborative dialogue (Swain, 1998, 2000) (collaborative rehearsal) which necessitates the role of language (both L1 and L2) as a cognitive tool in problem-solving and constructing knowledge and the importance of learner agency(socio-cultural). This suggests the value of viewing task-based language teaching and learning through the lens of “theoretical pluralism” (Ellis, 2008). No single perspective, be it cognitive, psycholinguistic, or socio-cultural, was sufficient by itself to reveal the interplay of tasks, setting, materials, teachers and learners in the data analysed in this research, and how the interplay shaped language learning opportunities through tasks (also see Batstone, 2010; Ellis, 2008, 2012; Mackey & Polio, 2009; Swain & Deters, 2007). In fact, the theoretical perspectives are complementary rather than in opposition.

10.6 Limitations and future research directions

The research has a number of limitations which warrant acknowledgement and provide areas for future research directions.

Firstly, the research was conducted at a leading high school in Vietnam which may have been differentiated from other schools in terms of student motivation to learn

and teacher expertise. However, even granting a limit to the generalisations that this point prompts, if we are seeking effective ways of language teaching, then Williams (1999) argues “surely it makes sense to look for ... best examples of language teaching rather than typical ones” (p.619).

Secondly, the present research can only reveal short-term learning through oral tasks in a rehearsal-performance condition (uptake within five minutes (Phase 1) and 15 minutes (Phase 2) of rehearsal). The evidence of uptake may be the result of short-term memory, under the pressure of a public performance. Future studies could use delayed public performance (without learners being aware of this) or delayed task repetition (Bitchener, 2004; Bygate, 1996, 2001; Bygate & Samuda, 2005) to trace longer-term learning.

Thirdly, the current research did not incorporate a control group (e.g., without (public) performance) because performance was an essential part of the task lessons in this teaching context. Although the results show that learning was pushed through a rehearsal-performance approach and that learners said public performance made them attend to form to convey their intended messages, it still remains unclear whether uptake in performance was due to the motivation of the public performance in prospect alone or the ‘linguaging’ (LREs) that had occurred in the rehearsal, or practice effect (public performance as the repeated performance). Future studies could investigate whether prior knowledge of an impending public performance (Skehan & Foster, 1997) would affect the incidence of LREs and how they are resolved in rehearsal and corresponding uptake in performance.

Fourthly, the present study only demonstrated the uptake by the interlocutors who asked for assistance or received feedback or carried out self-searches, or self-corrected, because the pattern in the data was that it was those learners who took up the items in the performance. The data did not provide a means of measuring learning for *both* dyad members. Research (e.g., van Lier, 1996) has shown learners benefit from assisting their peers and that learning also occurs for participating learners who do not initiate the LREs. It was also possible that students simply did not use, in the public performance, certain ideas discussed during rehearsal, thus obviating the use of

relevant L2 words/phrases targeted in LREs. It would be useful if future studies could use a combination of uptake and LRE-specific post-test items administered to both dyad members.

Fifthly, the current research used *dyadic* public performance, since this was the teachers' preferred way to engage students. It would be interesting to see whether individual presentation would bring about different results. An earlier experimental study (Skehan et al., 2012) suggested that expectation of an individual vs. pair transcription as a post-task activity had differential effects on the actual task performance.

Sixthly, the research used a small sample of data: only one task for each task type (convergent/divergent) and 24 dyads, eight for each proficiency group (HH, HL, LL). A larger sample which includes more tasks of the same type, and more dyads in each proficiency group would be useful in future research to achieve greater statistical strength for some variables.

Seventhly,³⁷ in the present research, I categorised LREs broadly into lexical and grammatical LREs. Categorising LREs in this way has also been quite popular in previous research (e.g., Kim & McDonough, 2008; McDonough & Sunitham, 2009; Swain & Lapkin, 2001; Williams, 1999, 2001). However, this broad categorisation may fail to capture how learners attended to different aspects of grammar such as tenses, subject-verb agreement or different lexical features such as verbs, concrete and abstract nouns and collocations. As a consequence, evidence on grammatical and lexical uptake in performance may not reflect the possible differential learning of these different aspects of grammar and vocabulary. Research (Laufer, 2005; Nation, 2001) has shown that instruction may be more or less beneficial, depending on whether lexical features are concrete or abstract nouns, single words or collocations. Likewise, empirical evidence has also indicated some grammatical features may be more amenable to instruction than others (Gass et al., 2005; Jeon, 2007). Hence future studies may consider using more detailed categorisation of LREs in order to investigate

³⁷ This limitation drew on Suzuki's (2009, pp.137-138) discussion of a need for a fine-grained categorisation of written LREs.

whether different lexical and grammatical features yield differential uptake in performance or not.

Finally, it would also be useful to further investigate how the length of LREs (see Ellis et al., 2001a, 2001b; Fortune, 2005; Fortune & Thorp, 2001; Loewen, 2005; Storch, 2008) and LRE strategies (e.g., whether learners request help from others, correct others or self-correct) (see Williams, 1999) have an influence on uptake in performance.

Despite these limitations, the original contributions of the thesis lie in its central focus on tasks in classrooms from both teaching and learning perspectives. The research has brought together not only task design features and task implementation (pre-task, rehearsal, performance, post-task) in action, but also teacher (and student) thinking and student learning (LREs and uptake). It has provided new insights into (public) performance and rehearsal for that performance from real world classrooms. It has supplied the first evidence on the effects of tasks and proficiency on the take-up in dyadic performance of LRE-specific language items addressed in rehearsal of the same task.

10.7 Contextual reflection

Previous research has shown that meaningful communication in Asian EFL contexts in general and in Vietnamese EFL classrooms in particular is rare (e.g., Le & Barnard, 2009; Tomlinson & Bao, 2004). The current thesis has shown a different picture. Within the context of a new task-based curriculum, the teachers developed and selected both task design features and implementation procedures (e.g., rehearsal-performance) that would better engage students in meaning-making in the target English. Students were active in meaningful communication through the tasks and in seeking L2 resources from each other to express their communicative intentions. The LREs, and LRE resolutions in rehearsal and uptake in performance provide robust evidence that learners focussed on form and language use during their task talk. Students also said they valued creativity in task performance. As a result, perhaps it is time to treat, with great care, stereotypes of Asian teachers as sole authorities and learners as passive receivers of knowledge (see Butler, 2011). The Vietnamese EFL classroom has been described as “a cultural island” characterised by teacher dominance and learner

passiveness (Le, 2001, pp.35-36, also see Chapter 1). This image might be a result of many factors such as the classroom context, the teacher, students and classroom tasks. But they are not typical of the classrooms observed in this study. The results of this study lend support to Littlewood's (2000) claim that

If Asian students do indeed adopt the passive classroom attitudes that are often claimed, this is more likely to be a consequence of the educational contexts that have been or are now provided for them, than of any inherent dispositions of the students themselves. (p.33)

In sum, the results provide encouragement for teachers not to 'buy into' stereotypes of Asian teachers and learners as not-to-be questioned knowledge providers and 'obedient' receivers of knowledge respectively. By considering task design and implementation choices that could engage students in task work (see 10.3), teachers can and do create beneficial learning opportunities for learners to actively participate in and learn through tasks.

10.8 Personal reflection

As stated in the introduction chapter, the thesis has personal value to me as an EFL high school teacher and now as a researcher. Conducting this research has helped me see the interdependence between research and pedagogy. The pedagogical implications addressed in 10.3 will be all directly relevant to me when I resume my job as a teacher of English. I would like to conclude this personal reflection with a saying I learnt by heart as a high school student: "All theory, dear friend, is gray, but the golden tree of life springs ever green" (Johann Wolfgang von Goethe),³⁸ not as a discouragement, but an *inspiration*. An inspiration to address the ever abiding need to bring 'live' data from real language classrooms so as to bridge the gap between theory and practice, between task research and task pedagogy. Through this thesis, I have realised such an inspiration.

³⁸ Taken from: <http://www.brainyquote.com/quotes/quotes/j/johannwolf150617.html>.

REFERENCES

- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 29-52). New York: Oxford University Press.
- Adams, R., & Newton, J. (2009). TBLT in Asia: Constraints and Opportunities. *Asian Journal of English Language Teaching*, 9, 1-17.
- Albert, A. (2011). When individual differences come into play: The effect of learner creativity on simple and complex task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 239-265). Amsterdam: John Benjamins.
- Alcón, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL Classroom. *International Journal of English Studies*, 7(2), 41-60.
- Alcón, E. (2009). Focus on form, learner uptake and lexical gains in learners' oral production. *IRAL*, 47(3-4), 347-365.
- Alcón, E., & García Mayo, M. P. (2008). Focus on form and learning outcomes in the foreign language classroom. In J. Philp, R. Oliver & A. Mackey (Eds.), *Child's play? Second language acquisition and the younger learner* (pp. 173-192). Amsterdam: John Benjamins.
- Alegría de la Colina, A., & García Mayo, M. P. (2007). Attention to form across collaborative tasks by low-proficiency learners in an EFL setting. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 91-116). Clevedon: Multilingual Matters.
- Alegría de la Colina, A., & García Mayo, M. P. (2009). Oral interaction in task-based EFL learning: The use of the L1 as a cognitive tool. *IRAL*, 47(3-4), 325-345.
- Alley, D. C. (2005). A study of Spanish II high school students' discourse during group work. *Foreign Language Annals*, 38(2), 250-258.
- Allwright, D. (1984). Why don't learners learn what teachers teach? The interaction hypothesis. In D. Singleton & D. Littlewood (Eds.), *Language learning in formal and informal contexts* (pp. 3-18). Dublin, Ireland: IRAAL.
- Allwright, D. (1996). Social and pedagogical pressures in the language classroom: The role of socialization. In H. Coleman (Ed.), *Society and the language classroom* (pp. 209-228). Cambridge: Cambridge University Press.
- Allwright, D. (2005). From teaching points to learning opportunities and beyond. *TESOL Quarterly*, 39(1), 9-31.

- Andon, N., & Eckerth, J. (2009). Chacun à son gout? Task-based L2 pedagogy from the teacher's point of view. *International Journal of Applied Linguistics*, 19(3), 286-310.
- Antón, M., & DiCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *The Canadian Modern Language Journal*, 54(3), 314-342.
- Appel, G., & Lantolf, J. P. (1994). Speaking as mediation: A study of L1 and L2 text recall tasks. *The Modern Language Journal*, 78(4), 437-452.
- Aston, G. (1986). Troubleshooting interaction with learners: The more the merrier? *Applied Linguistics*, 7(2), 128-143.
- Baecher, L. H., & Dang, T. B. T. (2011). Supervision of instruction in English as a foreign language: A Vietnamese perspective. *TESOL Journal*, 5, 3-20.
- Barnard, R., & Nguyen, G. V. (2010). Task-based language teaching (TBLT): A Vietnamese case study using narrative frames to elicit teachers' beliefs. *Language Education in Asia*, 1, 77- 86.
- Batstone, R. (2002). Contexts of engagement: A discourse perspective on 'intake' and 'pushed output'. *System*, 30(1), 1-14.
- Batstone, R. (2005). Planning as discourse activity: A sociocognitive view. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 277-295). Amsterdam: John Benjamins.
- Batstone, R. (2010). Issues and options in sociocognition. In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 3-23). New York: Oxford University Press.
- Berben, M., Van den Branden, K., & Van Gorp, K. (2007). "We'll see what happens": Tasks on paper and tasks in a multilingual classroom. In K. Van den Branden, K. Van Gorp & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 32-67). Newcastle: Cambridge Scholars Publishing.
- Bitchener, J. (2004). The relationship between the negotiation of meaning and language learning: A longitudinal study. *Language Awareness*, 13(2), 81-95.
- Bock, G. (2000). Difficulties in implementing communicative theory in Vietnam. *Teachers Edition*, 2, 24-30.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109.
- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. London: Continuum.

- Borg, S. (2009). Language teacher cognition. In A. Burns & J. C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 163-171). Cambridge: Cambridge University Press.
- Breen, M. (2009). Learners contributions to task design. In K. Van den Branden, M. Bygate & J. Norris (Eds.), *Task-based language teaching: A reader*. Amsterdam: John Benjamins.
- Breen, M., & Candlin, C. N. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics*, 1(2), 89-112.
- Breen, M., Hird, B., Milton, M., Oliver, R., & Thwait, A. (2001). Making sense of language teaching: Teachers' principles and classroom practices. *Applied Linguistics*, 22(4), 470-501.
- Brooks-Lewis, K. A. (2009). Adult learners' perceptions of the incorporation of their L1 in foreign language teaching and learning. *Applied Linguistics*, 30(2), 216-235.
- Brooks, F. B., & Donato, R. (1994). Vygotskian approaches to understanding foreign language learner discourse during communicative tasks. *Hyspania*, 77(2), 262-274.
- Brooks, F. B., Donato, R., & McGlone, J. V. (1997). When are they going to say "It" right? Understanding learner talk during pair-work activity. *Foreign Language Annals*, 30(4), 524 - 541.
- Brooks, L., & Swain, M. (2009). Languaging in collaborative writing: Creation of and response to expertise. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honour of Susan M. Gass* (pp. 58-89). New York: Routledge.
- Bruton, A., & Samuda, V. (1980). Learner and teacher roles in the treatment of oral error in group work. *RELC Journal*, 1, 49-63.
- Bryman, A. (2008). *Social research methods* (3rd ed.). Oxford: Oxford University Press.
- Burrows, C. (2008). An evaluation of task-based learning (TBL) in the Japanese classroom. *English Today*, 96, 11-16.
- Butler, Y. G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics*, 31(1), 36-57.
- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-146). Oxford: Macmillan.
- Bygate, M. (1999a). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, 3(3), 185-214.
- Bygate, M. (1999b). Task as context for the framing, reframing and un-framing of language. *System*, 7(1), 33-48.

- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 23-48). Harlow: Longman.
- Bygate, M. (2005). Oral second language abilities as expertise. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 104-127). New York: Palgrave MacMillan.
- Bygate, M. (2011). *Tasks in classrooms: Developing TBLT as a researched pedagogy: Invited Colloquia*. Paper presented at the 4th Biennial International TBLT Conference, The University of Auckland.
- Bygate, M., Norris, J., & Van den Branden, K. (2009). Understanding TBLT at the interface between research and pedagogy. In K. Van den Branden, M. Bygate & J. M. Norris (Eds.), *Task-based language teaching. A reader* (pp. 495-499). Amsterdam: John Benjamins.
- Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37-74). Amsterdam: John Benjamins.
- Bygate, M., & Samuda, V. (2009). Creating pressure in task pedagogy: The joint roles of field, purpose, and engagement within the interaction approach. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 90-116). New York: Routledge.
- Bygate, M., Skehan, P., & Swain, M. (2001). Introduction. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 1-20). Harlow: Longman.
- Candlin, C. N. (2001). Afterword: Taking the curriculum to task. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 229-243). Harlow: Longman.
- Carless, D. (2002). Implementing task-based learning with young learners. *ELT Journal*, 56(4).
- Carless, D. (2003). Factors in the implementation of task-based language teaching in primary schools. *System*, 31(4), 485-500.
- Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System*, 35(4), 595-608.
- Carless, D. (2008). Student use of the mother tongue in the task-based classroom. *ELT Journal*, 62(4), 331-338.
- Centeno-Cortés, B., & Jiménez, A. F. (2004). Problem-solving tasks in a foreign language: The importance of the L1 in private verbal thinking. *International Journal of Applied Linguistics*, 14(1), 7-35.

- Chung, I. F., & Huang, Y. C. (2009). The implementation of communicative language teaching: An investigation of students' viewpoints. *The Asia-Pacific Education Researcher*, 18(1), 67-78.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2 ed.). Hillsdale, NJ: Erlbaum.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). New York: Routledge.
- Cook, V. (2001). Using the first language in the classroom. *The Canadian Modern Language Review*, 57(3), 402-423.
- Cortazzi, M., & Jin, L. (1996). Cultures of learning: Language classrooms in China. In H. Coleman (Ed.), *Society and the language classroom* (pp. 169-206). Cambridge: Cambridge University Press.
- Coughlan, P., & Duff, P. (1994). Same task, different activities: Analysis of SLA task from an activity theory perspective. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 173-193). Westport, NJ: ALEX.
- Crabbe, D. (2003). The quality of language learning opportunities. *TESOL quarterly*, 37(1), 9-34.
- Crabbe, D. (2007). Learning opportunities: Adding learning value to tasks. *ELT Journal*, 61(2), 117-125.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: SAGE.
- Crookes, G. (1989). Planning and interlanguage variability. *Studies in Second Language Acquisition*, 11, 367- 383.
- Davis, R. (1997). Group work is NOT busy work: Maximizing success of group work in the L2 classroom. *Foreign Language Annals*, 30(2), 265-279.
- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46(3), 529-555.
- De Guerrero, M. (1994). Form and function of inner speech in adult second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language learning* (pp. 83-116). Westport, NJ: ALEX.
- De Guerrero, M. (1999). Inner speech as mental rehearsal: The case of advanced L2 learners. *Issues in Applied Linguistics*, 10(1), 27-55.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar'. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge: Cambridge University Press.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford: Blackwell.
- DeKeyser, R. (2007). Introduction: Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). Cambridge: Cambridge University Press.
- de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition*, 24(1), 81-112.
- Deng, C., & Carless, D. (2009). The communicativeness of activities in a task-based innovation in Guangdong, China. *Asian Journal of English Language Teaching*, 19, 113-134.
- DiCamilla, F., & Antón, M. (1997). The function of repetition in the collaborative discourse of L2 learners: A Vygotskian perspective. *The Canadian Modern Language Journal Review*, 53(4), 609-633.
- Dobao, A. F. (2012). Collaborative dialogue in learner-learner and learner-native speaker interaction. *Applied Linguistics*, 33(3), 229-256.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33-56). Westport, NJ: ALEX.
- Donato, R., & McCormic, D. (1994). A sociocultural perspective on language learning strategies: The role of mediation. *The Modern Language Journal*, 78(4), 453-464.
- Dörnyei, Z. (2001). *Motivational strategies in the classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences and instructed language learning*. Amsterdam: John Benjamins.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4(3), 275-300.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation and enhancement. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Malden, MA: Blackwell.
- Duff, P. A. (1986). Another look at interlanguage talk: Taking task to task. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 147-181). Cambridge: Newbury House.

- East, M. (2012). *Task-based language teaching from the teachers' perspective*. Amsterdam: John Benjamins.
- Eckerth, J. (2008). Investigating consciousness-raising tasks: Pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics*, 8(2), 121-145.
- Eckerth, J., & Siekmann, S. (Eds.). (2008). *Task-based language learning and teaching: Theoretical, methodological, and pedagogical perspectives*. Frankfurt: Peter Lang.
- Edwards, C., & Willis, J. (Eds.). (2004). *Teachers exploring tasks in English language teaching*. New York: Palgrave MacMillan.
- Eguchi, M., & Eguchi, K. (2006). The limited effect of TBL on EFL learners: A case study of English magazine projects. *Asian EFL journal*, 8(3), 207-225.
- Ellis, G. (1996). How culturally appropriate is the communicative approach? *ELT Journal*, 50(3), 213-218.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9(1), 1-19.
- Ellis, R. (1995). Uptake as language awareness. *Language Awareness*, 4(3), 147-160.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.
- Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193-220.
- Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, 51(Suppl. 1), 1-46.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (Ed.). (2005). *Planning and task performance in a second language*. Amsterdam: John Benjamins.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2009a). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221-246.
- Ellis, R. (2009b). The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474-509.
- Ellis, R. (2009c). SLA and Teacher Education. In A. Burns & J. C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 135-143). Cambridge: Cambridge University Press.
- Ellis, R. (2010a). Second language acquisition research and language-teaching materials. In N. Harwood (Ed.), *English language teaching materials: Theory and Practice* (pp. 33-57). New York: Cambridge University Press.
- Ellis, R. (2010b). Second language acquisition, teacher education and language pedagogy. *Language Teaching*, 43(2), 182-201.

- Ellis, R. (2012). *Language teaching research and language pedagogy*. MA: Wiley-Blackwell.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001a). Learner uptake in communicative ESL lessons. *Language Learning*, 51(2), 281-318.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001b). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, 35(3), 407-432.
- Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System*, 30(4), 419-432.
- Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21(2), 285-301.
- Ellis, R., Heimbach, R., Tanaka, Y., & Yamazaki, A. (1999). Modified input and the acquisition of word meaning by children and adults. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 63-114). Amsterdam: John Benjamins.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28(2), 339-368.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London: The Falmer Press.
- Farrokhi, F., & Gholami, J. (2007). Reactive and preemptive language-related episodes and uptake in an EFL class. *Asian EFL Journal*, 9(2), 58-92.
- Færch, C., & Kasper, G. (1986). The role of comprehension in second language learning. *Applied Linguistics*, 7(3), 257-274.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: SAGE.
- Flowerdew, J., & Miller, L. (1995). On the notion of culture in L2 lectures. *TESOL Quarterly*, 29(2), 345-373.
- Fontana, A., & Frey, J. H. (2005). The interview: From neutral stance to political involvement. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 695-727). Thousand Oaks: SAGE.
- Fortune, A. (2005). Learners' use of metalanguage in collaborative form-focused L2 output tasks. *Language Awareness*, 14(1), 21-38.
- Fortune, A., & Thorp, D. (2001). Knotted and entangled: New light on the Identification, classification and value of language-related episodes in collaborative output tasks. *Language Awareness*, 10(2-3), 143-160.
- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19(1), 1-23.
- Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second language classrooms. *Applied Linguistics*, 26(3), 402-430.

- Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18(3), 299-324.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 215-247.
- Fuji, A., & Mackey, A. (2009). Interactional feedback in learner learner interaction in task-based EFL classrooms. *IRAL*, 43(3-4), 267-301.
- García Mayo, M. P. (2002). The effectiveness of two form-focused tasks in advanced EFL pedagogy. *International Journal of Applied Linguistics*, 12(2), 156-175.
- García Mayo, M. P., & Pica, T. (2000a). Interaction among proficient learners: Are input, feedback and output needs addressed in a foreign language context? *Studia Linguistica*, 54(2), 272-279.
- García Mayo, M. P., & Pica, T. (2000b). L2 learner interaction in a foreign language setting: Are learning needs addressed? *IRAL*, 38(1), 35-58.
- Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, NJ: Erlbaum.
- Gass, S., & Alvarez Torres, M. (2005). Attention when? An investigation of the ordering effect of input and interaction. *Studies in Second Language Acquisition*, 27(1), 1-31.
- Gass, S., & Lewis, K. (2007). Perceptions about interactional feedback: Differences between heritage language learners and non-heritage language learners. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 79-99). Oxford: Oxford University Press.
- Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.
- Gass, S., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175-199). Mahwah, NJ: Erlbaum.
- Gass, S., Mackey, A., Alvarez-Torres, M., & Fernandez-Garcia, M. (1999). The effect of task repetition on linguistic output. *Language Learning*, 49(4), 549-581.
- Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, 55(4), 575-611.
- Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning*, 53(3), 497-545.
- Gass, S., & Varonis, E. (1986). Sex differences in NNS/NNS interactions. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*. Rowley: Newbury House.
- Gass, S., & Varonis, E. (1989). Incorporated repairs in nonnative discourse. In M. Eisenstein (Ed.), *The dynamic Interlanguage* (pp. 71-86). New York: Plenum Press.

- Gibbons, P. (2006). *Bridging discourses in the ESL classroom*. London: Continuum.
- Gilabert, G. (2007). The simultaneous manipulation of task complexity along planning time and (+/- Here -and -Now): Effects on L2 production. In M. D. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44-68). Clevedon: Multilingual Matters.
- Gilabert, G., Barón, J., & Ilanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *IRAL*, 47(3-4), 367-395.
- Givon, T. (1985). Function, structure and language acquisition. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 1008-1025). Hillsdale, NJ: Erlbaum.
- Givon, T. (2009). *The genesis of syntactic complexity*. Amsterdam: John Benjamins.
- Guariento, W., & Morley, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal*, 55(4), 347-353.
- Guk, L., & Kellogg, D. (2007). The ZPD and whole class teaching: Teacher-led and student-led interactional mediation of tasks. *Language Teaching Research*, 11(3), 281-299.
- Gutiérrez, A. G. (2008). Microgenesis, method and object: A study of collaborative activity in a Spanish as a foreign language classroom. *Applied Linguistics*, 29(1), 120-148.
- Han, Z. H. (2002). Rethinking the role of corrective feedback in communicative language teaching. *RELC Journal*, 33(1), 1-34.
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459-479.
- Hashemi, M. R. (2012). Reflections on mixing methods in applied linguistics research. *Applied Linguistics*, 33(2), 206-212.
- Ho, K. S. T. (2011). *An investigation of intercultural teaching and learning in tertiary EFL classrooms in Vietnam*. Unpublished PhD thesis, Victoria University of Wellington, Wellington, New Zealand.
- Hoang, V. V., Hoang, T. X. H., Dao, N. L., Vu, T. L., Do, T. M., & Nguyen, Q. T. (2006). *English 11*. Hanoi: Education Publishing House.
- Hoang, V. V., Hoang, T. X. H., Vu, T. L., Dao, N. L., Do, T. M., & Nguyen, Q. T. (2007). *English 12: Teacher's book*. Hanoi: Education Publishing House.
- Holliday, A. (1997). Six lessons: Cultural continuity in communicative language teaching. *Language Teaching Research*, 1(3), 212-238.
- Holliday, A. (2007). *Doing and writing qualitative research*. London: SAGE.

- Hood, M., Elwood, J., & Falout, J. (2009). Student attitudes toward task-based language teaching at Japanese universities. *Asian Journal of English Language Teaching*, 19, 19-47.
- Iwashita, N. (2001). The effect of learner proficiency on interactional moves and modified output in nonnative-nonnative interaction in Japanese as a foreign language. *System*, 29(2), 267-287.
- Iwashita, N., MacNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design. *Language learning*, 51(3), 401-436.
- Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24(2), 168-196.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34(2), 239-278.
- Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21(3), 421-452.
- Jeon, I. J., & Hahn, J. (2006). Exploring EFL teachers' perceptions of task-based language teaching: A case study of Korean secondary school classroom practice. *Asian EFL Journal*, 8(3), 192-206.
- Jeon, K. S. (2007). Interaction-driven L2 learning: Characterizing linguistic development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 379-403). Oxford: Oxford University Press.
- Kasanga, L. A. (1996). Effect of gender of the rate of interaction. Some implications for second language acquisition and classroom practice. *ITL. Review of Applied Linguistics*, 111(2), 155-192.
- Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with and high intermediate proficiency. In R. Ellis (Ed.), *Planning and task-based performance in a second language* (pp. 143-164). Amsterdam: John Benjamins.
- Kim, J., & Han, Z. (2007). Recasts in communicative EFL classes: Do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 269-297). Oxford: Oxford University Press.
- Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92(1), 114-130.

- Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System*, 37(2), 254-268.
- Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12(2), 211-234.
- Kim, Y., & McDonough, K. (2011). Using pre-task modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183-199.
- Kim, Y., & Petraki, E. (2009). Students' and teachers' use of and attitudes to L1 in the EFL classroom. *The Asian EFL Journal Quarterly*, 11(4), 58-89.
- King, N., & Horrocks, C. (2010). *Interviews in qualitative research*. Los Angeles: SAGE.
- Klapper, J. (2003). Taking communication to task? A critical review of recent trends in language teaching. *Language Learning Journal*, 27(1), 33-42.
- Kline, R. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, 50(2), 343-384.
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language awareness*, 3(2), 73-93.
- Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote metalanguage awareness in the French immersion classroom. In R. Johnson & M. Swain (Eds.), *Immersion education: International perspectives* (pp. 284-309). Cambridge: Cambridge University Press.
- Kramsch, C., & Sullivan, P. (1996). Appropriate pedagogy. *ELT Journal*, 50(3), 199-212.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. D. (1984). Immersion: Why it works and what it has taught us. *Language and Society*, 12(1), 61-64.
- Krashen, S. D. (1985). *The Input Hypothesis*. New York: Longman.
- Kuiken, F., & Vedder, I. (2005). Noticing the role of interaction in promoting language learning. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 353-381). Berlin: Mouton de Gruyter.

- Kuiken, F., & Vedder, I. (2012). Speaking and writing tasks and their effects on second language performance. In S. M. Gass & A. Mackey (Eds.), *The Routledge Handbook of second language acquisition* (pp. 364-377). London: Routledge.
- Kumaravadivelu, B. (1994). The Postmethod condition:(E)merging strategies for second/foreign language teaching. *TESOL Quarterly*, 28(1), 27-48.
- Kumaravadivelu, B. (2001). Toward post-method pedagogy. *TESOL Quarterly*, 35(4), 537-560.
- Kumaravadivelu, B. (2006a). *Understanding language teaching: From method to post method*. Mahwah, NJ: Erlbaum.
- Kumaravadivelu, B. (2006b). TESOL methods: Changing tracks, challenging trends. *TESOL Quarterly*, 40(1), 59-81.
- Kumaravadivelu, B. (2007). Learner perception of learning tasks. In K. Van den Branden, K. Van Gorp & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective*. Newcastle: Cambridge Scholars Publishing.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. California: SAGE.
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Los Angeles: SAGE.
- Lamb, T. E. (2000). Finding a voice: Learner autonomy and teacher education in an urban context. In B. Sinclair, I. McGrath & T. Lamb (Eds.), *Learner autonomy, teacher autonomy: Future directions* (pp. 118-127). Harlow: Longman.
- Lambert, C. P., & Engler, S. (2007). Information distribution and goal orientation in second language task design. In M. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 27-43). Clevedon: Multilingual Matters.
- Lantolf, J. P. (2000). Second language learning as a mediated process. *Language Teaching*, 33(2), 79-96.
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(1), 67-109.
- Lantolf, J. P. (2012). Sociocultural theory: A dialectical approach to L2 research. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 57-72). London: Routledge.
- Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 201-224). London: Erlbaum.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5(1), 223-250.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Le, V. C. (2001). Language and Vietnamese pedagogical contexts. *Teacher's Edition*, November, 34-40.
- Le, V. C., & Barnard, G. (2009). Curricular innovation behind closed classroom doors: A Vietnamese case study. *Prospect*, 24(2), 20-33.
- Lee, S. M. (2005). The pros and cons of task-based instruction in elementary classes. *English Teaching*, 60(2), 185-205.
- Leeser, M. J. (2004). Learner proficiency and focus on form in collaborative dialogue. *Language Teaching Research*, 8(1), 55-81.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Li, D. (1998). "It's always more difficult than you plan and imagine": Teachers' perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly*, 32(4), 677-703.
- Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177-196). Cambridge: Cambridge University Press.
- Littlewood, W. (2000). Do Asian students really want to listen and obey? *ELT Journal*, 54(1), 31-36.
- Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal*, 58(4), 319-326.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40(3), 243-249.
- Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, 54(1), 153-188.
- Loewen, S. (2005). Incidental focus on form and second language learning. *SSLA*, 27(3), 361-386.
- Loewen, S. (2006). Autonomy and language learning behaviour: The role of student initiation in meaning-focused L2 lessons In T. Lamb & H. Reinders (Eds.), *Supporting independent language learning: Issues and Interventions* (pp. 37-51). Frankfurt and Main: Peter Lang.
- Loewen, S. (2007). The prior and subsequent use of forms targeted in incidental focus on form. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 101-116). Oxford: Oxford University Press.

- Loewen, S., & Philp, J. (2012). Instructed second language acquisition. In A. Mackey & S. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 53-73). Sussex: Wiley-Blackwell.
- Long, M. H. (1983). Does second language instruction make a difference? A review of research. *TESOL Quarterly*, 17(3), 359-382.
- Long, M. H. (1985a). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77-99). Clevedon: Multilingual Matters.
- Long, M. H. (1985b). Input and second language acquisition theory. In S. Gass & S. Madden (Eds.), *Input and second language acquisition*. Rowley, Mass: Newbury House.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of research on second language acquisition* (pp. 413-468). New York: Academic.
- Long, M. H. (2005). A rationale for needs analysis and needs analysis research. In M. H. Long (Ed.), *Second language need analysis* (pp. 1-16). Cambridge: Cambridge University Press.
- Long, M. H. (2007). *Problems in SLA*. London: Erlbaum.
- Long, M. H. (2009). Methodological principles for language teaching. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 373-394). Sussex: Wiley-Blackwell.
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26(1), 27-56.
- Long, M. H., & Norris, J. (2009). Task-based teaching and assessment. In K. Van den Branden, M. Bygate & J. Norris (Eds.), *Task-based language teaching: A reader* (pp. 135-142). Amsterdam: John Benjamins.
- Long, M. H., & Porter, P. (1985). Groupwork, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19(2), 207-227.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Lynch, T., & Maclean, J. (2001). 'A case of exercising': Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching*

- pedagogic tasks: Second language learning, teaching and testing* (pp. 141-162). Harlow: Longman.
- Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48(2), 183-218.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37-66.
- Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *SSLA*, 21(4), 557-587.
- Mackey, A. (Ed.). (2007). *Input, interaction and corrective feedback in L2 learning*. Oxford: Oxford University Press.
- Mackey, A. (2012). *Input, interaction and corrective feedback in L2 learning*. Oxford: Oxford University Press.
- Mackey, A., Abbuhl, R., & Gass, S. M. (2012). Interactionist approach. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition*. London: Routledge.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *SSLA*, 22(4), 471-497.
- Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. NJ: Erlbaum.
- Mackey, A., Kanganas, A., & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly* 41(2), 285-312.
- Mackey, A., McDonough, K., Fuji, A., & Tatsumi, T. (2001). Investigating learners' reports about the L2 classroom. *IRAL*, 39(4), 285-307.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82(3), 338-356.
- Mackey, A., & Polio, C. (2009). Introduction. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 1-10). New York: Routledge.
- Malmqvist, A. (2005). How Does group discussion in reconstruction tasks affect written language output? *Language Awareness*, 14(2-3), 128-141.
- Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research*. Thousand Oaks: SAGE.
- McDonough, J., & McDonough, S. (1997). *Research methods for English language teachers*. London: Arnold.
- McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System*, 32(2), 207-224.

- McDonough, K., & Chaikitmongkol, W. (2007). Teachers' and learners' reactions to a task-based EFL course in Thailand. *TESOL Quarterly*, 41(1), 107-132.
- McDonough, K., & Sunitham, W. (2009). Collaborative dialogue between Thai EFL learners during self-access computer activities. *TESOL Quarterly*, 43(2), 231-254.
- McMillan, J. H., & Schumacher, S. (1993). *Research in education: A conceptual introduction*. New York: Harper Collins College Publishers.
- Meddings, L., & Thornbury, S. (2003). What dogme feels like. *Humanizing language teaching*, Year 5(6). Retrieved from <http://www.hltmag.co.uk/nov03/sart1.htm>.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20(1), 83-108.
- Michel, M. C. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (Vol. 2, pp. 141-173). Amsterdam: John Benjamins.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook of new methods*. Newbury Park, CA: SAGE.
- Mitchell, R., & Myles, F. (2004). *Second language learning theories* (2nd ed.). New York: Arnold.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12(1), 11-37.
- MOET. (2010). *Guide to implementation of standardized knowledge and skills: The English language subject for Vietnamese high school level*. Hanoi: Education Publishing House.
- Murphy, K. R., & Myers, B. (2004). *Statistical power analysis*. Mahwah, NJ: Erlbaum.
- Nakahama, Y., Tyler, D., & Van Lier, L. (2001). Negotiation of meaning in conversational and Information gap activities: A comparative discourse analysis. *TESOL Quarterly*, 35(3), 377- 405.
- Nassaji, H. (2010). The occurrence and effectiveness of spontaneous focus on form in adult ESL classrooms. *The Canadian Modern Language Review*, 66(6), 907-933.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.
- Newton, J. (2001). Options for vocabulary learning through communication tasks. *ELT Journal* 55(1), 30-37.

- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17(2), 164-187.
- Newton, J., & Kennedy, P. (1996). Effects of communication tasks on the grammatical relations marked by second language learners. *System*, 24(3), 309-322.
- Nguyen, B. (2003). *The design and use of English language teaching materials in Vietnamese secondary schools*. Unpublished PhD thesis, Victoria University of Wellington, Wellington, New Zealand.
- Nguyen, B. T. T., Newton, J., & Crabbe, D. (to appear). Preparing for Tasks in Vietnamese EFL High School Classrooms: Teachers in Action. In M. Thomas (Ed.), *TBLT in Asia: Challenges and Opportunities* (publisher TBA).
- Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal*, 47(3), 203-210.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL Quarterly*, 25(2), 279-295.
- Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.
- Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, 37(4), 589-613.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. MA: Heinle, Cengage Learning.
- Ohta, A. S. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Socio-cultural theory and second language learning* (pp. 51-78). Oxford: Oxford University Press.
- Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. London: Erlbaum.
- Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *The Modern language journal*, 86(1), 97-111.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. New York: Cambridge University Press.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21(1), 109-148.

- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 78-209). Amsterdam: John Benjamins.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36(4), 573-595.
- Park, S. (2010). The influence of pretask instructions and pretask planning on focus on form during Korean EFL task-based interaction *Language Teaching Research*, 14(1), 9-26.
- Pennycook, A. (1998). *English and the discourses of colonialism*. London: Routledge.
- Pham, H. H. (2007). Communicative language teaching: Unity within diversity. *ELT Journal* 61(3), 193-201.
- Phan, L. H. (2004). University classrooms in Vietnam: Contesting the stereotypes. *ELT Journal*, 58(1), 50-57.
- Phan, L. H. (2008). *Teaching English as an international language: Identity, resistance and negotiation*. Clevedon: Multilingual Matters.
- Phan, T. T. H. (2009). Impacts of Vietnam's social context on learners' attitudes towards foreign languages and English language learning: Implications for teaching and learning. *Asian EFL Journal*, 11(4), 169-188.
- Philp, J., & Mackey, A. (2010). Interaction research: What can socially informed approaches offer to cognitivists (and vice versa)? In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 210-227). Oxford: Oxford University Press.
- Philp, J., & Tognini, R. (2009). Language acquisition in foreign language contexts and the differential benefits of interaction. *IRAL*, 47(3-4), 245-266.
- Philp, J., Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: What factors foster a focus on form? *Language Awareness*, 19(4), 261-279.
- Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, outcomes? *Language Learning*, 44(3), 493-527.
- Pica, T. (1997). Second language teaching and research relationships: A North American view. *Language Teaching Research*, 1(1), 48-72.
- Pica, T. (2005). Second language acquisition research and applied linguistics. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 263-280). Mahwah, NJ: Erlbaum.

- Pica, T., Holliday, N., Lewis, E., Berducci, D., & Newman, D. (1991). Language learning through interaction: What role does gender play? *Studies in Second Language Acquisition*, 13(3), 343-376.
- Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9-34). Clevedon: Multilingual Matters.
- Pica, T., Lincoln-Porter, F., Paninos, D., & Linnell, J. (1996). Language learners' interaction: How does it address the input, output, and feedback needs of L2 Learners? *TESOL Quarterly*, 30(1), 59-83.
- Pinter, A. (2005). Task repetition with 10-year old children. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113-126). New York: Palgrave MacMillan.
- Pinter, A. (2007). What children say: Benefits of task repetition. In K. Van den Branden, K. Van Gorp & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 131-158). Newcastle: Cambridge Scholars Publishing.
- Plough, I., & Gass, S. M. (1993). Interlocutor and task familiarity: Effects on interactional structure. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 35-56). Clevedon: Multilingual Matters.
- Poole, A. (2005). The kinds of forms learners attend to during focus on form instruction: A description of an advanced ESL writing class. *Asian EFL Journal*, 7(3), 75-87.
- Porter, P. (1986). How learners talk to each other: Input and interaction in task-centered discussion. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 200-222). Rowley, MA: Newbury House.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Prabhu, N. S. (1992). The dynamics of the language lesson. *TESOL Quarterly*, 26(2), 225-241.
- Punch, K. F. (2006). *Developing effective research proposals* (2nd ed.). London: SAGE.
- Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies of Second language Acquisition*, 31(3), 437-470.
- Révész, A. (2012). Coding second language data validly and reliably. In A. Mackey & S. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 203-221). Sussex: Wiley-Blackwell.
- Révész, A., Sachs, R., & Mackey, A. (2011). Task complexity, uptake of recasts, and L2 development. In P. Robinson (Ed.), *Second language task complexity: Researching the*

- Cognition Hypothesis of language learning and performance* (Vol.2, pp. 203-235).
Amsterdam: John Benjamins.
- Richards, J. C. (2008). Second language teacher education today. *RELC*, 39(2), 158-177.
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*.
Cambridge: Cambridge University press.
- Richards, L. (2009). *Handling qualitative data: A practical guide*. London: SAGE.
- Richards, L., & Morse, J. M. (2007). *Readme first for a user's guide to qualitative methods* (2nd ed.). Thousand Oaks: SAGE.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27-57.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL*, 43(1), 1-32.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL*, 45(3), 193-213.
- Robinson, P. (2011a). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (Vol. 2, pp. 3-37). Amsterdam: John Benjamins.
- Robinson, P. (2011b). Task-based language learning: A review of issues. *Language Learning*, 61(Suppl. 1), 1-36.
- Robinson, P., & Gilabert, G. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL*, 45(3), 161-176.
- Ross-Feldman, L. (2007). Interaction in the L2 classroom: Does gender influence learning opportunities? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 53-77). Oxford: Oxford University Press.
- Samuda, V. (2001). Guiding the relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 119-140). Harlow: Longman.
- Samuda, V. (2007). *Tasks, design, and the architecture of pedagogic spaces*. Paper presented at the TBLT Conference 2007, University of Hawaii.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Hampshire: Palgrave Macmillan.

- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111-142). Amsterdam: John Benjamins.
- Savignon, S. J., & Wang, C. (2003). Communicative language teaching in EFL contexts: Learner attitudes and perceptions. *IRAL*, 41(3), 223-249.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 17-46.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Schostak, J. (2006). *Interviewing and representation in qualitative research*. Berkshire: Open University Press.
- Seedhouse, P. (1999). Task-based interaction. *ELT Journal*, 13(3), 149-156.
- Sheen. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361-392.
- Shehadeh, A. (2002). Comprehensible output, from occurrence to acquisition: An agenda for acquisitional research. *Language Learning*, 52(3), 597-647.
- Shehadeh, A. (2004). Modified output during task-based pair interaction and group interaction. *Journal of Applied Linguistics*, 1(3), 351-382.
- Silverman, D. (2010). *Doing qualitative research* (3rd ed.). London: SAGE.
- Sink, C. H., & Stroh, H. A. (2006). Practical significance: The use of effect sizes in school counseling research. *Professional School Counseling*, 9(5), 401-411.
- Retrieved from
<http://www33.homepage.villanova.edu/edward.fierros/practicalsignificance.html>.
- Skehan, P. (1996a). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1), 38-62.
- Skehan, P. (1996b). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 17-30). Oxford: Heinemann.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

- Skehan, P. (2002). A non-marginal role for tasks. *ELT Journal* 56(3), 239-295.
- Skehan, P. (2003). Task-based instruction. *Language Teaching Research*, 36(1), 1-14.
- Skehan, P. (2007a). Task research and language teaching: Reciprocal relationship. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in honour of Rod Ellis* (pp. 55-68). Oxford: Oxford University Press.
- Skehan, P. (2007b). Language instruction through tasks. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 289-301). New York: Springer.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3), 185-211.
- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University Press.
- Skehan, P., Xiaoyue, B., Qian, L., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, 16(2), 170-187.
- Slimani-Rolls, A. (1989). The role of topicalization in classroom language learning. *System*, 17(2), 223-224.
- Slimani-Rolls, A. (2005). Rethinking task-based language learning: What we learn from the learners. *Language Teaching Research*, 9(2), 195-218.
- Slobin, D. (2003). Language and thought online: Some cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the investigation of language and thought* (pp. 157-191). Cambridge: MA: The MIT Press.
- Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL QUARTERLY*, 39(1), 33-58.
- Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42(2), 181-207.
- Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 443-446). Thousand Oaks: SAGE.
- Storch, N. (2001a). Comparing ESL learners' attention to grammar on three different classroom tasks. *RELC Journal*, 32(2), 104-124.
- Storch, N. (2001b). How collaborative is pairwork? ESL tertiary students composing in pairs. *Language Teaching Research*, 5(1), 29-53.

- Storch, N. (2001c). *An investigation into the nature of pairwork in an ESL classroom and its effect on grammatical development*. Unpublished PhD thesis. The University of Melbourne, Australia.
- Storch, N. (2002a). Patterns of interaction in ESL pair work. *Language learning*, 52(1), 119-158.
- Storch, N. (2002b). Relationships formed in dyadic interaction and opportunity for learning. *International Journal of Educational Research*, 37(3-4), 305-322.
- Storch, N. (2005). Collaborative writing: Process, product and students' reflections. *Journal of Second Language Writing*, 14(3), 153-173.
- Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications for language development. *Language Awareness*, 17(2), 95-114.
- Storch, N., & Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research*, 14(4), 355-375.
- Storch, N., & Wigglesworth, G. (2003). Is there a role for the use of the L1 in an L2 setting? *TESOL Quarterly*, 37(4), 760-770.
- Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In D. P. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157-177). Clevedon: Multilingual Matters.
- Storch, N., & Wigglesworth, G. (2010). Students' engagement with feedback on writing: The role of learner agency/beliefs. In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 166-185). Oxford: Oxford University Press.
- Sullivan, P. N. (2000). Playfulness as mediation in communicative language teaching in a Vietnamese classroom. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 115-131). Oxford: Oxford University Press.
- Suzuki, W. (2009). *Languageing, direct correction and second language writing: Japanese university students of English*. Unpublished PhD thesis, University of Toronto, Toronto, Canada.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Journal Review*, 50, 158-164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.

- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogues. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.
- Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing*, 18(3), 275-302.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). Mahwah, NJ: Erlbaum.
- Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95-108). London: Continuum.
- Swain, M., Brooks, L., & Tocalli-Beller. (2002). Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, 22(1), 171-185.
- Swain, M., & Deters, P. (2007). "New" mainstream SLA theory: Expanded and enriched. *The Modern Language Journal*, 91(Focus issue), 820-836.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82(3), 320-337.
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: The uses of the first language. *Language Teaching Research*, 4(3), 251-274.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 99-118). Harlow: Longman.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37(3-4), 285-304.
- Swain, M., & Lapkin, S. (2006). "Oh, I get it now!" From production to comprehension in second language learning. In D. M. Brinton & O. Kagan (Eds.), *Heritage language acquisition: A new field emerging*. Mahwah, NJ: Erlbaum.
- Tomlinson, B., & Bao, D. (2004). The contributions of Vietnamese learners of English to ELT methodology. *Language Teaching Research*, 8(2), 199-222.

- Tarone, E., & Yule, G. (1989). *Focus on the language learner*. Oxford: Oxford University Press.
- Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effects of narrative type on learner output. *Language Learning*, 58(2), 439-473.
- Thomas, M. (2006). Research synthesis and historiography: The case of assessment of second language proficiency. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 279-298). Amsterdam: John Benjamins.
- Thornbury, S. (2012). Afterword: Dogme for beginners: The autonomy of the group. In K. Irie & A. Stewart (Eds.), *Realizing autonomy: Practice and reflection in language education contexts* (pp. 257-265). New York: Palgrave Macmillan.
- Truong, T. A., & Storch, N. (2007). Investigating group planning in preparation for oral presentations in an EFL class in Vietnam. *RELC Journal*, 38(1), 104-124.
- Van den Branden, K. (2006a). Introduction: Task-based language teaching in a nutshell. In K. Van den Branden (Ed.), *Task-based language education* (pp. 1-16). Cambridge: Cambridge University Press.
- Van den Branden, K. (2006b). Training teachers: Task-based as well? In K. Van den Branden (Ed.), *Task-based language education* (pp. 217-248). Cambridge: Cambridge University Press.
- Van den Branden, K. (Ed.). (2006c). *Task-based education*. Cambridge: Cambridge University Press.
- Van den Branden, K. (2009a). Mediating between pre-determined order and chaos. *International Journal of Applied Linguistics*, 19(3), 264-285.
- Van den Branden, K. (2009b). Diffusion and implementation of innovations. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 659-672). Sussex: Wiley-Blackwell.
- Van den Branden, K. (2012). Task-based language education. In A. Burns & J. C. Richards (Eds.), *The Cambridge guide to pedagogy and practice in second language teaching* (pp. 132-139). Cambridge: Cambridge University Press.
- Van den Branden, K., Bygate, M., & Norris, J. (2009). Task-based language teaching: Introducing the reader. In K. Van den Branden, M. Bygate & J. Norris (Eds.), *Task-based language teaching: A reader* (pp. 1-19). Amsterdam: John Benjamins.
- Van den Branden, K., Van Gorp, K., & Verhelst, M. (Eds.). (2007). *Tasks in action: Task-based language education from a classroom-based perspective*. Newcastle: Cambridge Scholars Publishing.
- Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy, and authenticity*. Harlow: Longman.

- Van Lier, L. (2004). *The ecology and semiotics of language learning*. Boston: Kluwer Academic Publishers.
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12(3), 287-301.
- VanPatten, B. (1996). *Input processing and grammar instruction*. New York: ALEX.
- VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-31). Mahwah, NJ: Erlbaum.
- Viera, A. J., & Garret, J. M. (2005). Understanding inter-observer agreement: The kappa statistic. *Family Medicine*, 37(5), 360-363.
- Villamil, O., & De Guerrero, M. (1996). Peer revision in the L2 classroom: Social cognitive activities, mediating strategies, and aspects of social behaviour. *Journal of Second Language Writing*, 5(1), 51-75.
- Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky. Volume 1. Thinking and speaking*. New York: Plenum Press.
- Watanabe, Y. (2008). Peer-peer interaction between L2 learners of different proficiency levels: Their interactions and reflections. *The Canadian Modern Language Journal*, 64(4), 605-635.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121-142.
- Watanabe, Y., & Swain, M. (2008). Perception of learner proficiency: Its impact on the interaction between an ESL learner and her higher and lower proficiency partners. *Language Awareness*, 17(2), 115-130.
- Weir, C., & Roberts, J. (1994). *Evaluation in ELT*. Oxford: Blackwell.
- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.
- White, R. (1988). *The ELT curriculum*. Oxford: Blackwell.
- Wickens, C. (2007). Attention to the second language. *IRAL*, 45(3), 177-191.
- Widdowson, H. G. (2003). *Defining Issues in English Language Teaching*. Oxford: Oxford University Press.
- Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language testing*, 14(1), 85-106.
- Wilkins, D. (1976). *Notional syllabuses*. Oxford: Oxford University Press.

- Williams, J. (1999). Learner-generated attention to form. *Language Learning*, 49(4), 583-625.
- Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, 29(3), 325-340.
- Williams, J. (2005). Form-focused instruction In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 671-712). London: Erlbaum.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Willis, J. (1996). *A framework for task-based learning*. Essex: Longman.
- Willis, J. (2004). Perspectives on task-based instruction: Understanding our practices, acknowledging different practitioners. In B. L. Leaver & J. Willis (Eds.), *Task-based instruction in foreign language education* (pp. 3-44). Washington, D.C: Georgetown University Press.
- Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making and classroom practice*. Cambridge: Cambridge University Press.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 24(1), 1-27.
- Yule, G., & Macdonald, D. (1990). Resolving referential conflicts in L2 Interaction: The effect of proficiency and interactive role. *Language Learning*, 40(4), 539-556.
- Zhang, E. Y. (2007). TBLT-Innovation in primary school English language teaching in Mainland China. In K. Van den Branden, K. Van Gorp & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 68-91). Newcastle: Cambridge Scholars Publishing.
- Zhao, S. Y., & Bitchener, J. (2007). Incidental focus on form in teacher-learner and learner-learner interactions. *System*, 35(4), 431-447.

APPENDICES

All the appendices are enclosed in a CD attached to this thesis.

Appendix 1 Ethics Approval

Appendix 2 Ethics for Phase 1

Appendix 2.1 Information sheets and consent forms for the school participant

Appendix 2.2 Information sheets and consent forms for teacher participants

Appendix 2.3 Information sheets and consent forms for student participants

Appendix 3 The textbook units under observation (Phase1)

Appendix 4 A sample of teacher interviews (Phase1)

Appendix 5 A sample of student interviews (Phase1)

Appendix 6 Ethics for Phase 2

Appendix 6.1 Information sheets and consent forms for the school participant

Appendix 6.2 Information sheets and consent forms for teacher participants

Appendix 6.3 Information sheets and consent forms for student participants

Appendix 7 The textbook units that provide contexts for Phase 2

Appendix 8 The tasks for the main data collection in Phase 2

Appendix 9 A sample of student interviews (Phase 2)