

THE EFFECT OF INPUT, REPETITION AND METACOGNITIVE AWARENESS ON
TASK PERFORMANCE IN SOUTH KOREAN EFL CLASSES

BY

SHAUN JUSTIN MANNING

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

Victoria University of Wellington

(2014)

ABSTRACT

The implementation of language learning tasks in EFL settings has been questioned for a number of reasons: limited use of L2, student overuse of a shared L1, a lack of connection between tasks and institutional demands, managing off-task behaviour and teacher unfamiliarity with tasks (Butler, 2011; Carless, 2004, 2008, 2009; Tinker Sachs, 2007). This study took an ecological approach (Van Lier, 2004) to investigate tasks and task implementation conditions designed to address the first four of these concerns. The study was a two-phase, mixed-method design in which I implemented ecologically valid tasks in an EFL course for South Korean university freshmen. The first phase was a one-semester exploratory study of six task-types with the aim of identifying deviations from the task-as-workplan. Differences from the workplan were discussed in relation to which stage of the task cycle, activity→ planning→ report (see J. Willis, 1996b) students re-interpreted and how these changes altered learning opportunities. The key findings were: (1) the students generally followed the workplan and engaged in extensive English-language collaborative assistance (Ohta, 2001). (2) Students rarely used the ‘planning’ phase of the task cycle to rehearse their public reports. They used this time to add new content, often using L1. (3) Throughout the task cycle, talk about grammar was rare except when one of the students was writing their team’s answer. (4) Echoing Carless (2008), students used markedly more L1 when the tasks were more “absorbing” because students worked to create interesting content rather than to develop their L2. (5) Of the six task types, the values clarification (VC) task was found to have best prompted students to discuss both lexis and grammar and to do so using relatively little L1.

Based on these findings, the second phase, a quasi-experimental intervention, was carried out to determine the effect of three task implementation conditions on VC task performance. The conditions were: (1) an input-processing activity (IPA) done either pre- or post-task (cf. J. Willis, 1996b); (2) task and procedural repetition (Kim, 2013); and (3) raising learners’ metacognitive awareness (MA) of the purposes of the task (cf. Ewald, 2004). The intervention took place across six weeks of regularly scheduled classes. Students were kept in the same small teams throughout the study. The main findings were: (1) Students did not use target items from the IPA in their task performances; (2) Neither task nor procedural repetition had a statistically significant impact on the number or type of LREs, the amount of L1 used or on the correct use of a targeted task-natural grammar form. However, LRE counts and L1 use showed unpredicted negative trends. (3) The MA

activity positively reversed the negative trends from repetition for some teams, indicating it provided a way to refocus participants on the learning objectives of the task particularly through the responses of other students.

The following conclusions were made. First, for students to use target language in task interaction, it should be essential to the task. Second, the inclusion of collaborative writing in speaking tasks promotes attention to grammar. The effects of task repetition are mitigated by familiarity with members of the team; and therefore, effects of task repetition reported elsewhere represent the interaction of language development plus a new interlocutor. Finally, the metacognitive awareness activity affected student output, but had no impact on the use of input, suggesting training in learning from input may be needed. This study concludes with suggestions for further research and language teaching pedagogy.

ACKNOWLEDGEMENTS

Although a thesis is the work of one person, it is the result of the efforts of many. I would like to take this opportunity to thank those who provided assistance, guidance and support along the way. First and foremost, this means expressing my deep and sincere gratitude to both of my supervisors, Dr. Jonathan Newton and Dr. Frank Boers of Victoria University of Wellington for their invaluable encouragement, support and critical review of this thesis. Their inspiration and guidance throughout the process has enabled me to express my intellectual curiosity and follow my research interests. I owe them a great debt.

In addition, I thank the three examiners, Dr. David Crabbe, Dr. Ineke Vedder, and Dr. Anne Feryok who have spent their valuable time, away from their own work, families and interests, to provide criticism and commentary on this work. Without critical review, research cannot progress, so their time and effort is greatly appreciated.

I am also indebted to many people from the research site. My colleagues, Gunther Breaux and Daniel Rhodes allowed me to record their voices for use as input in the study, and provided helpful feedback during the creation of the materials. They are both owed a heartfelt “Thank-you”. As is Dr. Jae-jin Sa, programme director for her kind permission and support for the project. Most of all, I thank the participating students whose energy, enthusiasm and curiosity are inspirational. Information from this study is intended to help them, and those who follow them, and they made it possible.

I also thank my fellow PhD students at Victoria University, particularly Robin Reid, Thi Bao Trang Nguyen, and Behnam Soltani for their support, encouragement, friendship and ideas. They listened patiently to my problems and complaints, and proved to be valuable sounding boards for ideas.

Finally, and most importantly, I thank my lovely wife, Kyunghee Back, for her kind-hearted, loving support and encouragement throughout this project and my teaching career. She was with me in the hard times and the good. All I do is for her.

TABLE OF CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	v
LIST OF TABLES	xiv
LIST OF EXTRACTS	xvi
LIST OF FIGURES	xviii
TRANSCRIPTION CONVENTIONS.....	xix
CHAPTER ONE: INTRODUCTION.....	1
1.1 Introduction.....	1
1.2 Introduction to tasks and task supported learning.....	3
1.2.1 Defining tasks	3
1.2.2 Tasks and language learning	4
1.2.3 Tasks and the syllabus.....	5
1.2.4 Researching tasks.....	5
1.3 Implementation of tasks in Asia.....	6
1.3.1 General issues in Asia	6
1.3.2 Specific issues in South Korea.....	6
1.4 Personal Statement.....	7
1.5 Motivation for and significance of the study	9
1.6 Summary	10
CHAPTER TWO: LITERATURE REVIEW	13
2.1 Tasks and language learning	13
2.1.1 Defining tasks	13
2.2 Theoretical underpinnings for tasks	15
2.2.1 Cognitive perspectives of task performance and learning	15
2.2.1.1 Input, output and noticing.....	15

2.2.1.2 Task Complexity	17
2.2.1.3 Summary of the cognitive perspective	19
2.2.2 Sociocultural perspectives of task performance and language learning	19
2.2.2.1 Engagement.....	20
2.2.2.2 Alignment	20
2.2.2.3 Assistance	21
2.2.2.4 Language Socialization	22
2.2.2.4 Summary of the sociocultural perspective	22
2.2.3 Ecological perspectives of task performance and learning.....	23
2.2.3.1 Ecology and affordances.....	23
2.2.3.2 Ecological research and ecological validity	24
2.2.4 This study's perspective	26
2.3 Implementing TBLT	27
2.3.1 Using tasks in the language classroom	27
2.3.1.1 Task-based, task-referenced, and task-supported approaches.....	27
2.3.1.2 Focused and unfocused tasks	28
2.4 Task implementation and task performance	29
2.4.1 Observing and interviewing: Teacher perceptions and behaviours	29
2.4.2 Experimental research: Complexity, Accuracy and Fluency (CAF).....	31
2.4.3 Language Related Episodes (LRE).....	31
2.4.4 L1 use in tasks.....	32
2.5 Willis' (1996) pedagogical framework for TBLT	35
2.5.1 The task cycle	35
2.5.1.1 The task stage.....	35
2.5.1.2 Planning and reporting stages	37
2.5.1.3 The pre- and post-task phases	38
2.5.2 The Willis' framework and syllabus design.....	39
2.6 Input and task performance	41
2.6.1 Input in the pre-task	41
2.6.2 Input in the post-task.....	44

2.7	Task repetition	44
2.8	Learner awareness and task-based learning	48
2.9	Chapter Summary	51
CHAPTER THREE: THE EXPLORATORY STUDY		53
3.1	Introduction	53
3.2	Purpose of the exploratory study	53
3.3	Methods	54
3.3.1	The research context	54
3.3.2	Participants	55
3.3.2.1	Student-participants	55
3.3.2.2	Human ethics	56
3.3.3	Data collection procedures	57
3.3.4	Data analysis procedures	58
3.3.5	Materials: The assigned textbook and supplementary tasks	59
3.3.5.1	The assigned textbook	59
3.3.5.2	Supplementary tasks	59
3.4	Results and discussion of the exploratory phase	63
3.4.1	Task performance	63
3.4.1.1	Failure to plan or rehearse the report	67
3.4.1.2	Failure to understand the input	72
3.4.2	Attention to grammar and meaning	73
3.4.2.1	Values clarification task	74
3.4.2.2	Listing tasks	76
3.4.3	L1 use	78
3.5	Summary of the exploratory study	79
CHAPTER FOUR: MAIN STUDY METHODS		81
4.1	Overview	81
4.1.1	Purpose statement of Phase 2	81
4.1.2	Research Questions	82

4.2	The Values Clarification (VC) Task	85
4.2.1	Describing the VC tasks used in the study.....	85
4.2.2	Doing the VC task in this study	86
4.3	The Participants.....	87
4.3.1	The Participating Students	87
4.3.2	Non-participating Students	88
4.3.3	The instructor and assisting colleagues.....	88
4.4	Materials.....	88
4.4.1	Tasks	88
4.4.2	Input-processing materials	90
4.4.3	Metacognitive awareness (MA) intervention materials.....	91
4.5	Task implementation conditions (independent variables).....	92
4.5.1	Pre-task vs. post-task input processing activity (IPA)	92
4.5.2	Task repetition (the original task vs. the repeated task).....	93
4.5.3	Metacognitive Awareness (MA) intervention	94
4.6	Overall research design	94
4.7	Data collection and analysis procedures	95
4.8	Dependent variables	97
4.8.1	Language Related Episodes (LREs)	97
4.8.2	L1 use: on and off topic	99
4.8.3	Use of a specified task-natural lexico-grammatical structure: <i>agree with</i> . 100	
4.9	Reliability	101
4.10	Chapter Summary	101
CHAPTER FIVE: VALUES CLARIFICATION TASK PERFORMANCE		103
5.1	Introduction	103
5.1.1	The data set	103
5.2	Performance of the task.....	104

5.2.1	The values clarification task-as-workplan	104
5.2.2	The values clarification task-in-process.....	104
5.2.2.1	Opinion-oriented approach: Team A5, Lesson 5 discussing prompt 1.....	105
5.2.2.2	Revision-oriented approach: Team D1, Lesson 5, discussing prompt 1	109
5.2.2.3	Summary of the two teams' task performance	113
5.3	Learning opportunities in the VC task	114
5.3.1	Learning through collaboration.....	114
5.3.2	Learning through mining the prompts	116
5.3.3	Learning of the course material	119
5.3.4	Summary of learning in the VC task.....	119
5.4	Student voices about learning from the VC task.....	120
5.4.1	The VC task and language development	121
5.4.1.1	The VC task and vocabulary development.....	121
5.4.1.2	The VC task and grammar development	122
5.4.1.3	The VC task and comprehension skills development.....	123
5.4.1.4	The VC task and speaking skill development	123
5.4.1.5	The VC task and learning through peer interaction.....	124
5.4.2	The VC task and content-knowledge development	124
5.5	Conclusions about the VC task	126
5.6	Chapter Summary.....	127
CHAPTER SIX: THE EFFECT OF AN INPUT PROCESSING ACTIVITY ON TASK PERFORMANCE.....		129
6.1	Introduction	129
6.1.1	Review of the research design	129
6.1.2	Research questions.....	130
6.2	Quantitative data and analysis.....	130
6.2.1	The effect of the input-processing activity on LREs	130
6.2.2	The effect of the input-processing activity on L1 use.....	133
6.2.3	The effect of the input processing activity on use of targeted items.....	138

6.2.4	The effect of the input processing activity on the correct use of <i>agree with</i>	140
6.2.5	Summary of quantitative input-condition data	143
6.3	A qualitative analysis of the role of pre-task input processing in task performance	144
6.3.1	Student observations and perceptions of the input	144
6.3.1.1	Authenticity.....	145
6.3.1.2	Obligatory use	146
6.3.1.3	Input order.....	150
6.3.1.4	L1 use.....	151
6.3.2	Classroom observations	156
6.3.3	Summary of the qualitative input-condition data.....	163
6.4	Discussion of the input-placement manipulation	164
6.4.1	Participants' previous experience with the tasks	165
6.4.2	Input-related factors	166
6.4.2.1	Modelling LREs in input.....	166
6.4.2.2	Factors preventing input being used	166
6.4.2.3	Incidental acquisition of a repeated lexico-grammatical structure.....	168
6.5	Chapter summary	168
CHAPTER SEVEN: THE EFFECT OF TWO TYPES OF TASK REPETITION ON TASK PERFORMANCE.....		171
7.1	Introduction	171
7.1.1	Review of the research design	171
7.1.2	Research questions.....	172
7.2	Quantitative results and analysis	172
7.2.1	The effect of procedural repetition.....	172
7.2.1.1	The effect of procedural repetition on language related episodes (LREs)	173
7.2.1.2	The effect of procedural repetition on L1-use.....	176
7.2.1.3	The effect of procedural repetition on 'AGREE+PREP' use.....	178
7.2.1.4	Summary of procedural repetition	178

7.2.2	The effect of task repetition	179
7.2.2.1	The effect of task repetition on LREs.....	179
7.2.2.2	The effect of task repetition on L1 use.....	180
7.2.2.3	Task repetition and ‘AGREE+PREP’.....	182
7.2.3	Summary of the effect of repetition on task performance	183
7.3	A qualitative perspective on procedural and task repetition	183
7.3.1	Student perspectives on procedural and task repetition.....	184
7.3.1.1	Student perspectives on procedural repetition.....	184
7.3.1.2	Student perspectives on task repetition	185
7.3.2	Student behaviour in task performance.....	188
7.3.2.1	Lack of engagement: not starting	188
7.3.2.2	Lack of engagement: speaking less	192
7.3.2.3	Task repetition as a ‘second chance to get it right’	195
7.4	Discussion of the task repetition results.....	198
7.4.1	Constant team membership or changing partners.....	199
7.5	Chapter summary	202
CHAPTER EIGHT: THE EFFECT OF A METACOGNITIVE AWARENESS (MA)		
ACTIVITY ON TASK PERFORMANCE		205
8.1	Introduction	205
8.1.1	Implementation of the MA task	206
8.1.2	Research questions.....	206
8.2	Quantitative data and analysis.....	207
8.2.1	The between-subjects effect of the MA task.....	207
8.2.1.1	The between-subjects effect of the MA task on LREs	208
8.2.1.2	The between-subjects effect of the MA task on L1 use.....	209
8.2.1.3	The between-subjects effect of the MA task on AGREE+PREP formation	210
8.2.1.4	Summary of between-subjects effect of the MA task.....	211
8.2.2	The within-subjects effect of the MA task.....	211
8.2.2.1	The within-subjects effect of the MA intervention on LREs	212
8.2.2.2	The within-subjects effect of the MA task on L1 use.....	215
8.2.2.3	The within-subjects effect of the MA task on AGREE+PREP formation.....	217

8.2.2.4 Summary of within-subjects effects.....	218
8.3 Case study analysis and discussion of MA intervention	219
8.3.1 Team A3	219
8.3.1.1 Task performance prior to the MA intervention	219
8.3.1.2 Talk during the MA intervention	220
8.3.1.3 Values clarification (VC) task performance after the MA intervention.....	223
8.3.1.4 Summary of Team A3.....	229
8.4 Chapter summary	229
CHAPTER NINE: CONCLUSIONS.....	233
9.1 Overview	233
9.2 Summary of the findings	234
9.2.1 Summary of the exploratory phase (Phase 1)	234
9.2.2 Summary of the intervention study using the values clarification task	235
9.2.2.1 The basis for the interventions	235
9.2.2.2 The intervention study	236
9.3 Limitations of the study.....	238
9.3.1 Small number of participants	238
9.3.2 Measure of L1 use.....	239
9.3.3 Performance, learning and use	239
9.3.4 The possibility of response bias	240
9.4 Theoretical implications	241
9.4.1 The role of input.....	241
9.4.2 The role of writing	242
9.4.3 The role of repetition	244
9.4.4 The role of metacognitive awareness.....	245
9.4.5 The approach to research	246
9.5 Pedagogical contributions	247
9.6 General Conclusions	249
9.7 Final personal statement.....	249

LIST OF REFERENCES	251
APPENDIX 1: SAMPLE VALUES CLARIFICATION TASK	271
APPENDIX 2: SAMPLE NATIVE-SPEAKER DIALOG AND LANGUAGE AWARENESS ACTIVITIES	273
APPENDIX 3: METACOGNITIVE AWARENESS RAISING ACTIVITY.....	276
APPENDIX 4: ETHICS SHEETS - ENGLISH.....	277
APPENDIX 5: STIMULATED RECALL INTERVIEW QUESTIONS	280
APPENDIX 6: ETHICS SHEET - KOREAN	281
APPENDIX 7: RECALL INTERVIEW QUESTIONS (KOREAN).....	283
APPENDIX 8: TASKS INVESTIGATED IN THE EXPLORATORY STUDY	284
APPENDIX 9: PHASE 2 DATA.....	288
APPENDIX 10: TARGET ITEMS FOR VC TASKS.....	294

LIST OF TABLES

Table 3.1. Composition of classes A and B.....	56
Table 3.2. Characteristics of the tasks used	61
Table 3.3. Schedule of tasks, topic and textbook chapters.....	62
Table 3.4. Summary table of observed task performance	64
Table 4.1. The Values Clarification task-as-workplan	87
Table 6.1. Arrangement of lesson, topic, task, and input-placement sequence.....	130
Table 6.2. The number of LREs per 10 minutes of task performance across input- placement condition.....	132
Table 6.3. Results of Mann-Whitney U tests for total LREs per 10min task talk across input condition.....	132
Table 6.4. The total number of turns using L1 per 10 minutes of task performance talk	135
Table 6.5. Results of Mann-Whitney tests for L1 use across Input Condition	135
Table 6.6. Results of Mann-Whitney tests for L1 across groups (Group 1 vs. Group 2)	136
Table 6.7. Results of Mann-Whitney U tests for Total L1 turns per 10 min.....	137
Table 6.8. L1 use for each team (expressed in number of turns per 10 minutes of talk).	138
Table 6.9. Correct and incorrect instances of AGREE+PREP (per 10 min.).....	141
Table 6.10. Results of Mann-Whitney U tests investigating AGREE+PREP use across Input Condition (task-then-input vs. input-then-task)	142
Table 7.1. Arrangement of topic and task repetition type	172
Table 7.2. The total number of LREs, L1-use turns and use of the targeted phrasal verb, “AGREE+PREP” across all teams	174
Table 7.3. RM-ANOVA results for LREs.	175
Table 7.4. LREs across task repetitions	180
Table 7.5. Total turns of L1-use per 10min VC task talk.....	181
Table 7.6. Results of Related-Samples Wilcoxon Signed Ranks tests for effects of task repetition on off-topic L1 use	181
Table 7.7. Results of Related-Samples Wilcoxon Signed Ranks tests for effects of task repetition on on-topic L1 use.....	182
Table 7.8. The number of correct, incorrect and total uses of ‘AGREE+PREP’(per 10 min)	182
Table 8.1. Between-subjects comparison of the number of LREs, turns of L1 and formation of AGREE+PREP across MA conditions	208

Table 8.2. Independent samples t-tests comparing LREs across MA condition	209
Table 8.3. Results of Mann-Whitney U tests comparing L1 use across MA condition ...	210
Table 8.4. Independent samples t-tests comparing LREs across MA condition	211
Table 8.5. Within-subjects comparison of Lesson 4 task performance with MA+ and MA- conditions from Lessons 5 and 6 on all variables.....	212
Table 8.6. RM-ANOVA results for LREs: total, grammatical and vocabulary	213
Table 8.7. Results of RM-ANOVA for LREs including only Lessons 4 – 6	214
Table 8.8. Paired samples t-tests comparing Lesson 4 vs. Lesson 5.....	215
Table 8.9. Paired samples t-tests comparing Lesson 4 vs. Lesson 6.....	215
Table 8.10. Number of turns of L1 talk for Lessons 4, 5 and 6 per MA condition.....	216
Table 8.11. Number of turns of L1 for Lessons 4, 5 and 6 per MA condition.....	217
Table 8.12. Related-Samples Friedman’s 2-Way ANOVA for correct formation of AGREE+PREP per MA condition.....	218
Table 8.13. Team A3 Total LREs, total L1 and L1-off topic use	219

LIST OF EXTRACTS

Extract 3.1. Planning in creative role play	67
Extract 3.2 Stephanie's presentation at Team 2	70
Extract 3.3. Doug's presentation at Team 5	70
Extract 3.4. Jigsaw task retelling	72
Extract 3.5. Laura's report to the class.....	73
Extract 3.6. LREg co-occurring with writing during a values clarification task	74
Extract 3.7. LREg co-occurring with writing in a listing task	76
Extract 4.1. LREg, selecting <i>have</i> or <i>has</i>	98
Extract 4.2. LREv, nominating the word <i>individual</i>	98
Extract 4.3. An LREv mediated with L1 (' <i>disregard</i> ').....	98
Extract 4.4. LREp about the pronunciation of 'wisdom'	99
Extract 5.1. Team A5, Lesson 5, discussion of prompt 1.....	106
Extract 5.2. Team D1, Lesson 5, discussing prompt 1	110
Extract 5.3. Team B3, Lesson 3, Helen, Sara and Sue, mining from the prompt	117
Extract 5.4. Jessie saying why the VC task is helpful for language development.....	121
Extract 5.5. Jessie claims that adjusting the sentence means using new words.....	122
Extract 5.6. Lucky commenting on writing in the VC task	122
Extract 5.7. Steve talking about the role of comprehension in the VC task	123
Extract 5.8. Sunny's statement on the relation of the VC task to school learning.....	124
Extract 5.9. Winnie on learning from more advanced peers.....	124
Extract 5.10. Jessie's view of the VC task as content knowledge consolidation.....	125
Extract 5.11. Steve's view that VC talk as a comprehension check of content	125
Extract 6.1. Nancy using the item ' <i>the range of</i> '	140
Extract 6.2. Lucky: native speaker input is helpful and complicated	145
Extract 6.3. Nada: native speakers use language students don't get exposed to.....	145
Extract 6.4. Jon: the input scripts were more realistic than textbooks.....	146
Extract 6.5. Mary's presentation of her group's opinion, 'learning is memorisation'	147
Extract 6.6. Jessie – testing beforehand will lead to use.....	148
Extract 6.7. Jessie's story of learning 'expert' during the test	149
Extract 6.8. Nada: writing new expressions into sentences	149
Extract 6.9. Jessie's preference for input-then-task	150
Extract 6.10. Janine on pretask input as source of ideas, not language	150

Extract 6.11. Lucky cannot remember the previous lesson	151
Extract 6.12 Sunny used L1 as a supplement	152
Extract 6.13. Winnie claimed a lack of English leads to L1 use.....	152
Extract 6.14. Jon talking about the relation of topic familiarity and L1 use.....	153
Extract 6.15. Lucky on the process of task talk in her group.....	153
Extract 6.16. Jessie on the depth of the conversation prompting L1 use.....	153
Extract 6.17. Sunny always wanted to give her opinion.....	154
Extract 6.18. Jessie – L1 use to fit in.	154
Extract 6.19. Nada – using L1 to ‘hang out’ and to ‘complete thoughts quickly’	155
Extract 6.20. Michelle – Korean appears immediately when talking to friends.....	156
Extract 6.21. Sara using the verb ‘agree’	168
Extract 7.1. Jon claims repeating the values clarification develops topical/content knowledge.....	184
Extract 7.2. Repeating the task is ‘practice makes perfect’	185
Extract 7.3. Jon – task repetition allows for creative enhancement.....	186
Extract 7.4. Lucky – task repetition is a chance to fix mistakes and add more	186
Extract 7.5. Nada and the missed opportunity to improve in the repeated lesson	187
Extract 7.6. Team B3, Lesson 1(VC ₁ -original)	188
Extract 7.7. Team B3, Lesson 3 (VC ₂ - original)	189
Extract 7.8. Group B3, Lesson 4 (VC ₂ -repeated)	190
Extract 7.9. Group B3, difficulty getting started, Lesson 5	191
Extract 7.10. Team A3, Lesson 3, (VC ₂ -original) task performance of the prompt	192
Extract 7.11. Team A3, Lesson 4, (VC ₂ -repeated) task performance of the prompt	194
Extract 7.12. B3, Lesson 3 (VC ₂ -original), resolving the LREv- ‘stereotype’ first attempt	195
Extract 7.13. Team B3, Lesson 4, (VC ₂ -repeated), resolving LREv- ‘stereotype’ second attempt	196
Extract 7.14. Janine wanted a variety of interlocutors.....	200
Extract 7.15. Jessie – the same group members increases L1 (Korean) use.....	200
Extract 8.1. Team A3, between MA+ and doing VC task in Lesson 5	223
Extract 8.2. Group A3’s L1 use in Lesson 5	224
Extract 8.3. Lesson 5, Team A3, working on prompt 2 after MA intervention	225
Extract 8.4. Group A3 spontaneous self-assessment in Lesson 5	228

LIST OF FIGURES

Figure 2.1: Overview of the TBL Framework (J. Willis, 1996a, p. 155).....	36
Figure 4.1. Sample Values Clarification Task, used in Lesson 3	85
Figure 4.2. The Metacognitive Awareness survey task questions.....	92
Figure 4.3. Comparison of activities in input-then-task and task-then-input conditions...	93
Figure 4.4. Overview of lesson, chapter and lesson condition for the study	95
Figure 5.1. Team A5's final version of prompt 1	109
Figure 5.2. Team D1's Final written version of prompt 1	112
Figure 5.3. Mary using gestures to organise task roles- she tells Eunice to write	113
Figure 5.4. Mary using gestures to indicate where to write.....	113
Figure 5.5. Team B3, Lesson 3 final version of prompt 1	118
Figure 6.1. The mean number and type of LREs by input condition.....	133
Figure 6.2. The number and type of L1 turns per input condition.....	136
Figure 6.3. Correctly and incorrectly formed AGREE+PREP combinations per input condition	143
Figure 6.4. Winnie holding the input sheets containing transcripts during pretask input	158
Figure 6.5. Teams A3 (front) and A4 (rear) facing the board during feedback on LA	158
Figure 6.6. Sunny holding the VC prompts during VC task talk.....	159
Figure 6.7. Sunny writing the answers.....	159
Figure 6.8. Team D1, looking at the pretask input transcript and LA	160
Figure 6.9. Teams D1 and D2 looking at the whiteboard during LA feedback	161
Figure 6.10. Team D1, VC task talk – ‘Mary: “ <i>I have experienced becoming stupid</i> ” ...	161
Figure 6.11. Mary suggests deleting ‘more and more stupid’	162
Figure 6.12. Mary dictating while Eunice writes and clarifies	163
Figure 7.1. Mean number of total, grammatical and vocabulary LREs per 10min	175
Figure 7.2. The number of turns of L1 use across lessons.....	177
Figure 7.3. Number of correct and incorrect uses of AGREE+PREP	178
Figure 7.4. Final written version of A3, Lesson 3, (VC ₂ -original) task performance	194
Figure 7.5. Final written version of A3, Lesson 4, (VC ₂ -repeated) task performance	194
Figure 8.1. Total LREs, LREg and LREv by lesson	213

TRANSCRIPTION CONVENTIONS

This is a list of how speech is represented in this thesis.

Symbol	Meaning
word	any English word that was spoken
CAPS	words or syllables spoken more prominently than surrounding speech
°word°	words between the degree sign indicate words spoken more softly than surrounding speech
?	rising intonation (not necessarily a question)
.	falling, final intonation (not necessarily a statement) or a brief pause
,	continuing intonation
...	a long pause
(3.0)	a pause measured in seconds
h	an outbreath (more h's means a longer breath)
*h	an in-breath
//	onset of overlapping speech
\\	end of overlapping speech
L1	indicates what follows in italics is Romanized Korean
<i>anyeonghaseyo</i>	words in italics are Romanized Korean
안녕하세요	Korean (L1) words written in Korean (Hangul) script follow the Romanization
(tr= hello)	translations of L1 into English follow the Korean script in parentheses, indicated
(comment)	comments appear in parentheses
he- hello	dashes indicate hesitations for single speakers or interruptions across speakers
(<i>unint</i>)	unintelligible speech

CHAPTER ONE: INTRODUCTION

1.1 Introduction

Interactive language learning tasks have a long history in the English as a Second Language (ESL) classroom, so much so that they appear to some researchers to be the current orthodoxy in ESL methodology (Andon & Eckerth, 2009; Littlewood, 2004). Typically, classroom tasks involve students communicating, most often by speaking, to other students who need the information they receive in order to complete the task satisfactorily (Ellis, 2003; Nunan, 2004; Samuda & Bygate, 2008; J. Willis & Willis, 2007).

Interactive tasks were originally based on the premise that gaps of information, between what learners are able to communicate and what they are unable to communicate, push learners to discuss language and adapt it so that it is comprehensible to their hearers (Ellis, 2003, 2006; Long, 1996, 2007; Skehan, 1998; J. Willis, 1996b). Through the interaction involved in completing the task, learners adjust their talk to their own level, discuss items they do not know, and, therefore focus on items they need, not items predetermined by a syllabus.

However, teachers in English as a Foreign Language (EFL) settings, particularly in Northeast Asia, have been hesitant to use tasks in their classrooms. The reasons include the limited use of English (L2) by students, student overuse of a shared first language (L1), a lack of connection between tasks and institutional demands, difficulty managing the students' off-task behaviour and teacher unfamiliarity with tasks (Butler, 2011; Carless, 2004, 2008, 2009; Tinker Sachs, 2007).

In addition, the use of tasks in EFL classroom has been questioned by teachers and researchers because: tasks are modified by teachers and students to suit the local context (e.g. Carless, 2004; Kumaravadivelu, 1991, 2007; Littlewood, 2007); and, because the task was brought into the class by a researcher for research purposes and was not an original part of the students' course of study (Butler, 2011). Few studies have been carried out using ecologically valid tasks – tasks that existed in that context prior to the research

and that were selected by the instructor for use with the students to meet their learning needs and not the researcher's agenda (Van Lier, 2004), although Ngyuen (2013) is a recent exception.

This study sought to address this gap by first evaluating the tasks used in a South Korean EFL classroom and then implementing the most suitable of them under various conditions to determine if task performance could be improved under these conditions. It followed an ecological approach to research (Van Lier, 2004, p. 193) as it only examined tasks and implementation conditions existing in that context in previous semesters. In this way, the students and researcher experienced the tasks as natural and logical elements of their EFL course (See Chapters 4 and 5, below).

The first implementation condition addressed the effect of using recordings of native speaker task performance as pre-task models for students to follow (Kim & McDonough, 2011) or as post-task sources of language for students to analyse (Hawkes, 2011; J. Willis, 1996b). There has been no study to determine which position, pre or post, will best aid subsequent task performance.

In addition, repeating a task has been shown to benefit learners, (e.g. Hawkes, 2011; Kim, 2013; Lynch & Maclean, 2000, 2001; Shintani, 2012) but most studies involve repeating the exact task with new partners, while little research has investigated the use of a task procedure with different topics several times in an EFL course. Kim (2013) is an exception but her study was 'parachuted' into a middle school class. Moreover, because the course textbook and instructors at the current research site often re-used a task procedure across lessons, an ecologically valid investigation of task performance over several procedural repetitions was warranted. Alongside this, an investigation into exact task repetition was carried out to confirm the ecological validity of the earlier research.

The final implementation condition was the use of a metacognitive awareness raising survey. Managing learner use of their first language and reducing off-task behaviour were both identified as issues for implementing tasks (Butler, 2011; Carless, 2004, 2008). Although learner strategy training has a rich tradition (e.g. Bialystok, 1990; Dörnyei & Scott, 1997) most involves explicit training (e.g. Lam, 2009; Naughton, 2006; Seifoori & Vahidi, 2012). In contrast, this study used an implicit, peer-driven approach to engage reflection on task performance, thereby providing insight into the role of peers in task performance in addition to the peer's use of language.

The research followed a mixed-methods design, and was made up of two consecutive studies, a qualitative exploratory study followed by a quasi-experimental quantitative intervention (Creswell, 2002; Hesse-Biber, 2010). It was informed by both cognitive and sociocultural approaches to language learning. The following theoretical constructs informed the study: the Output Hypothesis (Swain, 1985, 1995), the Interactionist Approach (Gass & Mackey, 2007; Mackey & Gass, 2006), Schmidt's noticing hypothesis (Schmidt, 1990); sociocognition and the notion of alignment (Atkinson, Churchill, Nishino, & Okada, 2007; Atkinson, 2002, 2010; Batstone, 2010, 2012); and sociocultural theory (Lantolf & Thorne, 2007; Lantolf, 2006), particularly task engagement (Platt & Brooks, 2002) and language socialisation (Duff & Kobayashi, 2010; Duff, 2010). Finally, the ecological approach (Van Lier, 2000, 2004) informed the overall approach to this research and efforts were made at all stages of the study to maintain the ecological validity of the design and implementation of this study (Bronfenbrenner, 1979; Van Lier, 2004). As such, the data analysis involves both statistical and interactional analyses (Ellis & Barkhuizen, 2005, pp. 163–196) and the conclusions are based on the integration of both.

The thesis is organised in the following way. Chapter 1 introduces the study. Chapter 2 outlines the theoretical influences that gave rise to the study. Chapter 3 presents the exploratory study. Chapter 4 details the research design and methods for the intervention study. Chapter 5, 6, 7 and 8 provide the results and discussion related to the values clarification task and each intervention in turn. Chapter 9 outlines the conclusions together with the study's limitations, implications and recommendations for research and pedagogy.

1.2 Introduction to tasks and task supported learning

1.2.1 Defining tasks

Of the many definitions of task (e.g. Breen, 1987; Ellis, 2003; Nunan, 2004, 2006; Samuda & Bygate, 2008; Skehan, 1998; J. Willis, 1996b), this study will use that of Samuda and Bygate (2008) because it updated Ellis (2003, p16). "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 product or process, or both” (Samuda & Bygate, 2008, p. 69).

1.2.2 Tasks and language learning

There are two main theoretical frameworks used to explain how tasks help language learning. These will be discussed in Chapter 2, but I briefly outline them here. Cognitive approaches claim that task interaction provides a chance to notice a gap between what the learner can and cannot do in the L2 and take steps to close the gap (Schmidt, 1990). The interaction in tasks provides learners with a chance to negotiate meaning, that is to give and receive interactional feedback to achieve a comprehensible level of input and output (Long, 1996). Cognitive approaches also claim that characteristics of tasks influence the language produced by learners by requiring them to allocate cognitive resources to different aspects of speech production (Ellis, 2005, 2009b; Gilabert, Baron, & Levkina, 2011; Robinson, Cadierno, & Shirai, 2009; Robinson, 2001, 2007, 2012; Skehan, Bei, Li, & Wang, 2012; Skehan, 1998, 2003, 2009).

Sociocultural approaches view learning as a dynamic process through which a learner moves from being unable to do a task without the help of others to being able to do it individually (Moll, 1990; Vygotsky, 1986). Tasks provide what Vygotsky (1986) called a Zone of Proximal Development (ZPD) which, in the field of second language acquisition (SLA) has been defined 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” (Ohta, 2001, p. 9). This means, if a learner accomplishes a task with assistance from a teacher or group members that they could not do without the assistance, the learner is in the process of learning, as the learner is now closer to being able to do it alone (Foster & Ohta, 2005, p. 415). It need not be the full task that a learner cannot do alone. It could be something smaller such as completing a sentence or adding the correct morpheme to indicate past tense. Task interaction provides opportunities for learners to provide assistance to each other as they perform the task thereby leading to learning (Ohta, 2001).

This study holds the view that both cognitive and sociocultural perspectives offer complementary perspectives that together lead to a fuller understanding of the process of language learning in classrooms. This view concurs with Van Lier’s claim that “both qualitative and quantitative methods of analysis can be used, depending on what it is that

is relevant to the research question” (Van Lier, 2004, p. 195).

1.2.3 Tasks and the syllabus

Samuda and Bygate (2008), extend distinctions made by others (Ellis, 2003; J. Willis, 1996b) of task-based and task-supported language learning. Samuda and Bygate (2008) categorise three main approaches to the use of tasks in language learning and teaching. These uses reflect different types of relationship between tasks and other elements of the syllabus. The most task-centric approach is task-based learning and teaching (TBLT) in which the syllabus is defined by tasks. These tasks are chosen for their ability to replicate real world activities the learners need to do in the target language. Language items are selected for teaching based on the needs of the learners. Finally, students are assessed based on their task performance (p. 58). The second most task-centric approach is task-referenced learning and teaching (TRLT), which differs from TBLT in that tasks are used for assessment but “makes no assumption about the extent to which tasks may be used in teaching” (p. 59). The least task-centric approach to the use of tasks is task-supported learning and teaching (TSLT). In TSLT, tasks are important elements in the programme, but are done with other types of activity, do not form the basis of the syllabus and may or may not be used in assessment (p.60). TSLT has also been referred to as a ‘weak form’ of TBLT (e.g. Carless, 2004).

1.2.4 Researching tasks

This brief introduction shows that there are multiple definitions of tasks, multiple theoretical accounts for how tasks work, and multiple ways of implementing tasks in classrooms. Likewise, there are multiple research approaches. First, experimental research has led to competing frameworks for task-based syllabus design (Robinson, 2007, 2010, 2012; Skehan et al., 2012; Skehan, 1998). The experimental approach seeks, among other things, to identify task design conditions that increase or decrease the difficulty of a task, thereby enabling a sequence of tasks for a syllabus to be created (e.g. Robinson, 2010). Second, a descriptive-interpretative approach, typically classroom-based research, has shed light on sociocultural factors that influence task performance. This approach seeks to describe and explain learner behaviour in task performance, under the assumption that task performance is learning (See, for example, Atkinson et al., 2007; Foster & Ohta, 2005; Lantolf & Thorne, 2007; Lantolf, 2006; Ohta, 2001; Philp, Walter, & Basturkmen, 2010; Platt & Brooks, 2002; Swain & Lapkin, 1998; Vygotsky, 1986). A

third, ecological, approach stems from sociocultural theory yet accepts both experimental and descriptive-interpretative research. Researchers working within an ecological approach examine language, participants and tasks as elements of a dynamic and interrelated system of relationships among learners, the classroom setting, and education system (Bronfenbrenner, 1979; Van Lier, 2004). The ecological approach requires study of learners in the context in which they learn. The current study is situated in the natural context of teaching and learning for the participants, and this ecological validity allows me to make claims about these learners' performance that other, non-ecological research may not be able to make.

1.3 Implementation of tasks in Asia

1.3.1 General issues in Asia

Butler (2011) recently provided an overview of how communicative language teaching (CLT) and task-based teaching (TBLT) have been implemented in school systems in Asia. She found that although many countries have mandated the use of communicative and/or task-based teaching there has been difficulty implementing them due to conceptual, classroom and societal constraints. Butler lists four unsolved issues and concerns: the implementation of tasks in exam-oriented educational contexts, how to teach grammar in TBLT, student use of L1, and top-down policy implementation (Butler, 2011, pp. 46–49).

This study addresses two of these. First it asks which tasks promote more focus to grammar during task interaction and why. Second, it examines how task implementation factors including the placement of an input processing activity (IPA) pre or post-task, task repetition and a metacognitive awareness activity influenced, among other things, L1 use.

1.3.2 Specific issues in South Korea

South Koreans hold English in high regard as a means for educational and social advancement and as a means to improve their nation through globalization and an export driven economy. It is also held in a negative light as being a means to maintain current social inequity through unequal access to English teaching resources (e.g. Jeon, 2009; Park, 2009; Song, 2011, 2012).

To address the issue of social inequity, starting in 1995, the Korean government implemented the following changes to improve English education: (1) Korean teachers of English, both in-service and pre-service were given additional training, (2) the starting age for English in the national curriculum was lowered to 3rd grade of elementary school, and (3) guarantees were made to ensure one native speaker in every middle and high school in the country (Jeon, 2009; J.-K. Park, 2009).

These initiatives have not produced the communication skills the government expected (Lee *et al.*, 2011). Teachers have been sticking to grammar-translation and audiolingual methods (Beaumont & Chang, 2011). New teachers who are trained in more recent teaching methods have been pressured by more senior teachers not to use them (Shin, 2012). The native speakers of English who were hired as assistant teachers were often not qualified teachers and unsure of their roles in the class vis-à-vis their Korean counterparts. This resulted in many cases where neither the Korean teacher-led nor the native speaker-led lessons contained much in the way of communication (S. Park & Manning, 2012). And finally, the college entrance examination in Korea, a multiple-choice reading and listening test, provides no incentive to teach more communicatively (Lee M.B. *et al.*, 2011; Jeon, 2009).

As a result, Korean students often reach university with the ability to do relatively well on multiple choice English tests, but many have never spoken or written their own ideas in English. For approximately ten per cent of the participants in this study, I was the first native speaker of English they had ever studied with, despite the government's efforts to the contrary. And for almost all of them, I was the first teacher to expect them to produce an idea or opinion of their own and to react to the ideas of others in class.

1.4 Personal Statement

I served both as researcher and classroom instructor. This means that the participants were

my students and my collaborators in all aspects of this research¹. Efforts to minimise the potential impact of such a relationship on the research findings will be discussed in greater detail later. This section instead will talk about my approach to teaching English through tasks and demonstrate how this interest led to this study.

I had been an instructor at the university where the research was conducted for 11 years, teaching English as a Foreign Language (EFL) in the ‘English for non-English majors’ programme, called “Practical English” in which the study took place. I also taught EFL to undergraduate English majors in the English College. In the English College I used tasks that followed the Willis’ framework for task based teaching (J. Willis & Willis, 2007; J. Willis, 1996a, 1996b). I found them to be well liked by my students and easy to implement.

In the Practical English department, however, conversations with other instructors in the programme revealed concerns about the effectiveness of using tasks. My colleagues were concerned about the students’ lower level of English proficiency – an issue raised by Willis (J. Willis & Willis, 2007; J. Willis, 1996b). They also raised arguments similar to ones raised by the teachers reported in Carless (2004): students speak in L1 (Korean) too much, (b) students who finish early get bored, (c) they don’t ‘learn’ they ‘practice’, and (d) the students need to pass standardized tests (in our case, for graduation).

In 2005, the university administrators required Practical English to be a sustained content language teaching (SCLT) program (Brinton, 2003; Murphy & Stoller, 2001; Pally, 2000) with the objective of preparing students for the English-medium instruction they would receive in their major areas of study. I felt that ‘academic tasks’ such as note-taking, dictogloss (Wajnryb, 1988), retelling and discussion would be useful tasks for the students to be able to do. Over the years, the range of tasks changed. The tasks investigated in the exploratory study were the tasks I had used previously as a teacher, not a teacher-researcher. So, I had a lot of experience using tasks, but had never really examined in detail what and how my students learned as they did them, or if some tasks

¹ This study received Human Ethics Research approval from Victoria University, RM#18762; the information sheets and informed consent forms, both English and Korean can be found in Appendixes 4, 5 and 6.

were better than others were at promoting the kinds of collaborative interaction known to promote learning. This was my personal motivation for the study.

1.5 Motivation for and significance of the study

The combination of my personal experience as an instructor along with some of the issues highlighted by researchers in the use of tasks in Asia provided the motivation for this study. It first aimed to investigate tasks used in an ecologically valid setting, not ones parachuted in for study by researchers. By choosing to use what teachers, in this case, my former colleagues and I, used for teaching purposes, this study tries to bridge the gap between research and practice in EFL.

One point to verify was the degree of similarity of the task-in-process, i.e. the task performance, to the task-as-workplan (Breen, 1987, 1989). And if they varied from the workplan, in what ways did the changes impact on the learning opportunities from the task (cf. Crabbe, 2003, 2007; Mori, 2002, 2004). Furthermore, if students took strategic or even L1 routes to finish the task, could interventions be developed that would lead them to perform the task closer to the workplan?

Many tasks researched so far have been either oral or written, but not integrated across modes, with dictogloss (Wajnryb, 1988) being a notable exception. In dictogloss tasks, students listen to a dictated passage and then work together in pairs or small groups to reconstruct it. If the passage is of a suitable level, students must share ideas about the reconstruction as none of them would have understood or remembered it perfectly. Hence, both speech and writing are required for successful dictogloss performance. Because Williams (2012) presents a strong argument for including writing in SLA teaching and research, and because nearly all the tasks I used when teaching integrated speech and writing to some degree, I felt that including ecologically valid, situated tasks which integrated speaking and writing would contribute to both pedagogical and theoretical knowledge.

Butler (2011) pointed out that one of the main issues with implementing tasks into contexts where multiple-choice tests were all important, such as Korea, was the relation

of TBLT to improving grammar and when to include grammar instruction in the task cycle. Willis proposed a task cycle with a focus on form at the end of it, after the main task (D. Willis & Willis, 1996; J. Willis & Willis, 2007; J. Willis, 1996a, 1996b). Despite its logical appeal based on Schmidt (1990), the effectiveness of examining and processing samples of language after the task has never been tested.

Another way for learners to improve language use is through manipulation of task conditions (See, for example Robinson, 2001; Skehan, 2009). Because both Robinson's and Skehan's approaches to sequencing a syllabus require a strong TBLT version which in many cases, including mine, cannot be implemented, I investigated other elements of task design that promote attention to grammatical accuracy while simultaneously focusing on meaning. One such element was repeating the task, which has been claimed in the literature claimed to be beneficial for language development (Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005, 2007; White, 2011)

As Butler (2011) and Carless (Carless, 2004, 2007) pointed out, teachers in monolingual classrooms complain that students use L1 to complete tasks. Carless (2008) added that more L1 was used in more engaging tasks, negating the learning opportunity through using L2. My own experience bore this out, so, I investigated if making learners metacognitively aware of the language learning function of tasks would reduce L1 use and change task performance.

This study, therefore, aims to contribute to the research literature on task performance by highlighting the interactive processes that learners undergo when doing tasks designed by their teacher for their learning needs in an ecologically valid setting. It also adds to our knowledge of how input, task repetition and metacognitive awareness raising affect task performance. It also provides instructors with applicable research on tasks, the findings of which they could implement in their EFL classrooms.

1.6 Summary

This chapter introduced the themes of the current study. It introduced some of the central issues about tasks and task-based learning and teaching. It sketched a rough outline of

theoretical approaches to tasks, the relation of tasks to language and content syllabus design and issues in implementation of tasks in Asian contexts. It then outlined my experience of teaching with tasks in Asia and how this mapped onto gaps in the research. The next chapter provides a detailed account of literature relevant to this study.

CHAPTER TWO: LITERATURE REVIEW

The previous chapter introduced the study, its ecological context and overall purpose. This chapter reviews the literature that informed it and highlights the relevant gaps in this literature that created the need for this study. It starts first by looking briefly at tasks and their role in language learning. Then it reviews pedagogical and theoretically motivated approaches to task based learning and teaching (TBLT) and the literature on implementation of TBLT. It then moves on to examine research on the role of input in task performance, the role of task repetition in language learning and the research on metacognitive awareness (MA), which, in this study, will be investigated as a way of improving task performance.

2.1 Tasks and language learning

2.1.1 Defining tasks

The multiple definitions of tasks is one of the main issues in researching tasks (Butler, 2011; Seedhouse, 2005; Swan, 2005) and in their implementation in classroom practice (e.g. Butler, 2011; Carless, 2004, 2007, 2009). Ellis (2003, pp. 4–5) lists nine definitions of task; Samuda and Bygate (2008, p. 63) list four, and Willis and Willis also list four (2007, p. 12). There is some overlap in these lists, but across three books about tasks there are twelve different definitions, not counting the ones each of the authors then create. The large number of definitions is a clear cause for concern in the field, so two of the more recent books on tasks refine and clarify earlier definitions.

First, Willis and Willis (2007, p.13) reformed Skehan's (1998, p. 96) definition into a list of six questions that can be asked to determine how "task-like" an activity is:

1. Does the activity engage learners' interest?
2. Is there a primary focus on meaning?
3. Is there an outcome?
4. Is success judged in terms of the outcome?

5. Is completion a priority?
6. Does the activity relate to real world activities?

This list provides a guide for teachers to follow when developing or using tasks in the language classroom. Although appearing to be yes-no questions, they allow for a range of answers along a cline of yes—no and the degree to which they are answered yes that makes a task more or less task-like.

In a similar way, based on and extending Ellis' (2003, pp. 9–10) criterial features of tasks, Samuda and Bygate offer what they call “a working definition of task,”:

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 (Samuda & Bygate, 2008, p. 69)

They claim that language use “is the key element of the definition” (p. 66). Another key point of this definition is the dual nature of tasks: they teach both content and procedure. The content (language) is necessary for the procedure to be carried out, that is, “language is both means and end” (Samuda & Bygate, 2008, p.69), but language development is the “overall aim” of the task. In order to enable language development, task designers may focus on using the final product of the task to push language development, or they may focus on using the process of doing the task as the vehicle of development, or they may use both.

In the latter definition, there is no mention of learners' interest or real world activity as in the Willis' (2007) set of questions. Likewise, the Willis' set of questions does not refer to the idea that learning can come from the product or the process of doing the task, and that completion or the process of completing can drive language learning. This does not mean that the Willis' ignore this element of tasks, rather, in her framework for task based learning Jane Willis explicitly separates the task from a public report about or public performance of the task (J. Willis, 1996a, 1996b) thereby acknowledging the process of task performance in language learning.

This study will adopt Samuda and Bygate's (2008) definition as ‘interest’ can be engaged without language use, through images or music, for example, and language use is a driving force behind learning (see the next sections). In addition, the relation to real world

activities may be less clear for some tasks, for example pedagogical tasks (see Ellis, 2003). However, it is recognised, that there is little separating the two approaches to task definition.

2.2 Theoretical underpinnings for tasks

There are two general theoretical approaches to language learning, one cognitive and the other socio-cultural. The next two sections will only discuss the elements of each perspective relevant to the current study as it is beyond the scope of this study to provide a comprehensive account of each approach. The reader is referred to Skehan (1998) and Lantolf (2000, 2006) for more complete treatment of these approaches.

2.2.1 Cognitive perspectives of task performance and learning

2.2.1.1 Input, output and noticing

Cognitive perspectives to task performance and language learning are generally framed in information processing theory, a conception of the mind in which mental activity is explained through the metaphor of the computer (Skehan, 1998). Terms commonly used in applied linguistics such as ‘input’ and ‘output’ originate in this metaphor (for example Krashen, 1982; Swain, 1985). Input refers to language that is read, or heard, by the learner and output refers to language spoken, or written, by the learner. Both input (Krashen, 1982) and output (Swain, 1985) have well-established roles in language learning.

Schmidt proposed the Noticing Hypothesis in which it was claimed that only language that was noticed, whether from input or from output, could be learned (Schmidt, 1990, 2001). Schmidt limited noticing to directing attention to “surface structure of utterances in the input – instances of language rather than any abstract rules or principles” (Schmidt, 2001, p. 5).

Noticing was included in Swain’s Output Hypothesis (Swain, 1985). Later developments of the hypothesis proposed three functions for output in language learning (Swain, 1995). The first function was to prompt learners to notice any gaps between what they wanted to

express and were able to express. The second was to allow learners the opportunity to test hypotheses about the target language and receive feedback on them. The final function of output was metatalk in which the learners reflected consciously on their output. All three of these functions require the learners to notice either their own language output (the first and third functions) or require other learners (or the teacher) to notice the speaker's language and provide feedback on it.

Long (1996) proposed that negotiation of meaning was necessary for interactive language learning. He claimed that learning in interaction was driven by a process of learners adjusting their talk to a mutually comprehensible level through three processes: clarification requests, confirmation checks and comprehension checks. For these processes to engage, one member of the interaction must notice a problem. In this way, he emphasises the role of the interlocutor in noticing – either through supplying the input or through reacting to the other's output.

To sum up, noticing is a central element to language learning from the cognitive perspective, yet planning tasks so that noticing of particular items occurs is problematic. How can a teacher pre-plan what a learner will attend to while interacting? There have been a variety of task-like approaches to induce noticing, for example, input processing approaches (VanPatten & Cadierno, 1993; VanPatten, 1996, 2002), consciousness raising (Fotos, 1994; Rutherford, 1987), and data driven learning (Johns, 1994), among others. From a direct task-based approach, pushing noticing has been done through manipulating features of task design (see Robinson, 2001, 2010; Skehan, 2003, 2009), through post-task consciousness raising activities implemented as part of the task cycle (D. Willis & Willis, 1996; J. Willis, 1996b), and through the use of focused tasks (Ellis, 2003; cf. Loschky & Bley-Vroman, 1993).

I turn now to a brief discussion of Skehan's and Robinson's approaches to task design, and later discuss Willis' task cycle (Section 2.5).

Following Schmidt's (1990) claim that: "Task demands are a powerful determinant of what is noticed ... The information committed to memory is essentially the information that must be heeded in order to carry out a task" (p. 143), Skehan created The Limited Attentional Capacity (LAC) model (Skehan et al., 2012; Skehan, 1998, 2001, 2003, 2009). This model explains how learners improve their language during task performance and provides criteria for how tasks should be sequenced. It is based on an information-

processing view of cognition, and claims that limits on working memory cause a trade-off between attention to meaning and attention to form. In the model, tasks are designed, classified, and ordered based on the following criteria:

1. Code complexity
 - Linguistic complexity and variety
 - Vocabulary load and variety
 - Redundancy and density
 2. Cognitive complexity
 - Cognitive familiarity*
 - Familiarity of topic and its predictability
 - Familiarity of discourse genre
 - Familiarity of task
 - Cognitive processing*
 - Information organisation
 - Amount of computation
 - Clarity and sufficiency of information given
 - Information type
 3. Communicative stress
 - Time limits and time pressure
 - Speed of presentation
 - Number of participants
 - Length of texts used
 - Type of response
 - Opportunities to control interaction
- (Skehan, 1998, p. 99)

This model proposes that increasing task complexity along any of these dimensions will increase the learners' attention to the content of the task and reduce the amount of attention available to address the accuracy and complexity of output. However, this classification does not give any indication of weighting or ordering of items within it (Kuiken & Vedder, 2007) thereby reducing the predictability and testability of the model (Robinson, 2007).

2.2.1.2 Task Complexity

Working from a cognitive perspective on tasks Robinson proposed the Triadic Componential Framework (TCF) for classifying tasks (Robinson et al., 2009; Robinson & Gilabert, 2007; Robinson, 2001, 2007, 2010, 2011a, 2011b). The TCF classifies tasks along three main categories: task complexity (TC), task condition (TCn), and task difficulty (TD). The first, task complexity relates to the cognitive factors involved in task performance. These are then divided into resource-dispersing and resource-directing

variables. The former make performative/procedural demands and the latter make cognitive/conceptual demands. Most of the research on the TCF has focused on task complexity, because Robinson (2011b) claims “in task-based syllabi, pedagogic tasks should be sequenced solely on the basis of increases in their cognitive complexity” (p.14). Task condition refers to interactive factors and the demands made on learners by participation variables and participant variables. The former refers to how the interaction is planned to take place, and the latter refers to the people in the interaction and their relationships. Task difficulty refers to learner variables, learner ability and affective variables relevant to the task performance.

Robinson extended the TCF to the SSARC model (simplify/stabilise – automatise – restructure/complexify) of syllabus design (Robinson, 2010). This model is based on two principles. First, “only the cognitive demands of tasks contributing to their intrinsic conceptual and cognitive processing complexity are sequenced” (p.247). Second, “increase resource-dispersing dimensions of complexity first, and then increase resource-directing dimensions” (*ibid.*). In this model of syllabus design, the first version of a task should be a simple/stabilising (SS) version. This first task should be simple on all dimensions of task complexity. In order for later versions to automatise (A) learners’ access to their interlanguage, TC should be increased on resource-dispersing dimensions, such as reducing planning time. The final version of the task should restructure and complexify (RC) the learners’ interlanguage. This is done by increasing TC on resource-directing dimensions such as, for example, increasing the number of elements or requiring intentional reasoning (p.252).

The SSARC model requires repeating a task (or task type) several times each one increasing TC along one or the other of the two dimensions. However, Robinson does not mention how many or how frequent such repetitions should be. Task repetition will be discussed in section 2.4, but, repetitions could be considered a kind of integrative planning (Bygate & Samuda, 2005). Since planning time is a resource-dispersing element of TC, even if we decrease or eliminate planning time in later versions of the task, having done the simpler versions may confound increases in TC because resource-dispersing elements would not necessarily be increasing in complexity if the tasks were similar in structure and topic. This leads to the same criticism of the TCF as was levelled at the LAC model, the relative weight given to each item in the TCF and interactions among the items need to be accounted for in order to arrange a syllabus and to date, they have not.

2.2.1.3 Summary of the cognitive perspective

Both Skehan and Robinson agree that the learning benefit of a task derives from using language to complete the task, and that the actual benefit is determined by the manner of the learner's allocation of cognitive resources to the task. However, Skehan and Robinson each offer different ideas of the nature of these resources and their allocation. As yet, this issue has not been fully resolved (see Robinson, 2011a).

In addition, although both Skehan and Robinson acknowledge the role of the interlocutor in tasks they stress that only cognitive elements should be included in task design and sequencing decisions (Robinson, 2007; Skehan, 2009). However, much of the impetus for noticing is theorised to originate in interaction breakdowns (Long, 1996), a gap in ability to express oneself, or a desire to test and get feedback on one's language (Swain, 1995). These will necessarily depend on who is doing the interacting, for example, their relative ages, genders and proficiency levels (Plough & Gass, 1993) or their willingness to communicate (Macintyre, Dörnyei, Clément, & Noels, 1998). These factors have origins both in the individual and in the individual's relation to other individuals in the social context. Therefore, although cognitive factors have been shown to be a useful point of origin for syllabus planning, they are not the only influence on classroom task performance and language learning. Influences on task performance that originate from interlocutors and the process of interaction have typically been researched from the sociocultural perspective, to which I now turn.

2.2.2 Sociocultural perspectives of task performance and language learning

An alternative approach to learning in tasks stems from a sociocultural view of learning. Originating in the work of Vygotsky (1986) this approach sees language learning as a mediated process of knowledge construction in which individual and social processes are interdependent. Sociocultural theory (SCT) holds that learning is mediated through tools, interaction or symbols (Vygotsky, 1986). Mediation can originate external to the learner, but the goal of learning is to allow the learner to mediate within him/herself. Vygotsky calls this the move from the interpersonal to the intrapersonal. Tasks, as interactive language learning activities, provide different forms of mediation for this move. Therefore, to study tasks from an SCT perspective, is to examine the participation of the learners and their use of language during the task, often from a microgenetic perspective rather than focusing on pre and post-tests, which typifies a cognitive approach. In SCT

approaches to language learning, participation is not only central to learning, but it is learning in action (see Lantolf, 2006).

As with the discussion of cognitive approaches, above, the purpose of this section is not to provide a complete account of SCT approaches to language learning, but to briefly introduce the aspects of SCT that directly relate to the present study: engagement, alignment, assistance, language socialisation and alignment.

2.2.2.1 Engagement

There are two types of engagement in TBLT: language engagement and task engagement (Platt & Brooks, 2002). Language engagement refers to learners' repetition, manipulation or planning of linguistic forms prior to using them (Ohta, 2001; Platt & Brooks, 2002). This form of engagement has been linked to language development (e.g. Swain & Lapkin, 2000).

Task engagement is conceived by Platt and Brooks (2002) as being more than doing the task. They argue task engagement involves "higher mental processes ... particularly those of volition (effort) and selective attention" (Platt & Brooks, 2002, p. 386). They demonstrate that task engagement can be shown at the verbal level, by more target language use, less procedural talk and increased fluency, and at the paralinguistic level, through posture, voice tone and reduction in non-functional behaviours (Platt & Brooks, 2002, p. 373). Task engagement, therefore, involves learners behaving in ways that promote language learning.

2.2.2.2 Alignment

A similar notion to engagement is alignment. Alignment is defined as the dynamic, coordinated interaction between learners and the target language along with any and all available affordances for learning (Atkinson et al., 2007, p. 171). Atkinson *et al* list the affordances (cf. Van Lier, 2000) for learning, of particular interest to this study are the following three:

- Tools: including learning materials (literacy tools), gaze and gesture (embodied tools), and interaction (social tools)
- Other people: the role and identity of interlocutors (e.g. peers and the instructor)

- Social situations and purposes: the social conventions involved creating and maintaining a functioning classroom; for example: instruction, conversations, and lessons

(Atkinson et al., 2007, p. 172)

Thus, the notion of alignment broadens the scope of investigation from learner and language to include how learners use all potential resources available to them while performing tasks.

Task engagement, above, could, then, be considered as learner behaviour that aligns with the task designer's intentions for performance of the task, i.e. the task-as-workplan. However, during task performance learners can, and do, align in ways unintended by the task designer or the instructor. The task gets completed but the learners' approaches to it differed greatly (cf. Breen, 1987, 1989 who reports similar issues while working from a more cognitive approach). This divergence in alignment indicates different learning opportunities for different learners. Examination of different patterns of alignment among learners performing the same task may lead to a better understanding of how learners and tasks interact to create learning opportunities.

2.2.2.3 Assistance

From a cognitive perspective, negotiation of meaning was proposed as a central process in language learning (Long, 1996). However this view was criticised because researchers found few instances of negotiation of meaning occurring in classroom data, negotiation of meaning is face threatening and disruptive to communication, it emphasises what was not successful, and is difficult to identify from transcripts (Foster & Ohta, 2005, p. 407).

From the SCT perspective, Ohta (2001) demonstrated that learners actively assisted each other when performing tasks. Assistance involved a greater range of behaviour than negotiation of meaning. It could be explicit, for example scaffolding (cf. Donato, 1994) or implicit, for example waiting for one's partner to finish. According to Ohta (2001) assistance offered a more complete perspective on how learners helped each other construct knowledge in interaction.

Foster and Ohta (2005) outlined four interactional processes through which learners provide explicit assistance to other learners as they interacted: co-construction, other

correction, self-correction and continuers; which they defined as follows:

- Co-construction: defined as “joint creation of an utterance”, involving one or more than one person chiming in to create the final utterance.
- Other-correction: a peer correcting their partner.
- Self-correction: a participant corrects their own utterance without a signal from another person.
- Continuers: words that express interest in what the other speaker is saying and encourage them to keep speaking (Foster & Ohta, 2005, pp. 420–1)

Assistance in the form of self- or other-correcting has been termed Language Related Episodes (LREs) (Swain & Lapkin, 1995, 1998). These are one of great interest in the current study and are examined in greater detail in Section 2.4.3.

2.2.2.4 Language Socialization

Duff and Kobayashi (2010) define language socialization as “The process by which novices or newcomers in a community or culture gain communicative competence, membership and legitimacy in the group by means of social interaction and (often) overt assistance, and by participating in the activities of the group” (p.78).

Although Duff and Kobayashi (2010) researched L2 learners in Canada and their socialization into Canadian university culture, the process of language socialisation is relevant to the current study of Korean learners in Korea. The learners were novices to the university (first year students), to the language classroom, and to the teaching method (TBLT) (see Chapter 1). They had to acquire a new set of behaviours in addition to L2 English in order to become full members of the classroom community and the acquisition process of these behaviours was mediated through language – from the instructor and their peers.

2.2.2.4 Summary of the sociocultural perspective

SCT is not a uniform approach, yet it has informed a large body of research into tasks contributing to our knowledge of how learners in language classrooms learn (for example Coughlan & Duff, 1994; Donato, 1994; Duff & Kobayashi, 2010; Duff, 2002; Lantolf & Thorne, 2007; Lantolf, 2006; Mori, 2002; Philp et al., 2010). The constructs briefly outlined here: engagement, alignment assistance and language socialisation are directly

relevant to the present study as they mediate learning outcomes by influencing the performance conditions.

2.2.3 Ecological perspectives of task performance and learning

Van Lier (2004) applied earlier ideas about ecology in the social sciences (Bronfenbrenner, 1979) to language learning. The ecological approach has much in common with SCT and could, therefore, be included in the previous section, but Van Lier makes distinctions between the two that are relevant to this study and therefore I have made a third section.

2.2.3.1 Ecology and affordances

The ecological perspective, metaphorically extends the natural science of ecology to the language classroom (Van Lier, 2004). This approach considers language-learning situations (e.g. classrooms) to be similar in important ways to ecological systems found in nature. In ecological systems, all organisms exist in a network of relations with all the other organisms, plants and minerals. Changes within the system or to the surrounding environment influence this relational network. Likewise, classrooms consist of a dynamic system of *relations* among active participants (for example, the power relationship between teacher and student). They are not static ‘context’ to be described. Nor are they ‘variables’ to be controlled for, or even eliminated through the use of out-of-classroom experiments. From an ecological perspective, context is the “heart of the matter” (Van Lier, 2004, p. 5).

Each communicative setting – each ecological context, contains a range of ‘things’ that aid or hinder communication– words, gestures, gaze, and social conventions, to name a few. Van Lier calls these *affordances*. More precisely defined; “an affordance ...[is] a relationship between an organism [the learner] and the environment [learning context] that signals an opportunity for or inhibition of action” (Van Lier, 2004, p. 4). While communicating, people select from the range of affordances available in their context to express themselves. A shopkeeper, for example, might use words, hold up fingers, or use the display on the cash register to show a customer the total price of purchased goods. In this example, language, gesture and tool are different affordances and meet the communicative needs of the shopkeeper in different ways. Each context has a “semiotic budget” of affordances similar to the energy budget of an ecosystem (Van Lier, 2000, p.

252). The decision to use a specific affordance depends on what the shopkeeper (the participant) needs to communicate, what affordances are available at that time, and which affordance is most useful to her or his needs at that instant.

Likewise, language-learning situations also have semiotic budgets of affordances. The difference between learning contexts and general contexts lies in the purpose of the communication, not the fact that it is situated. In a learning context, like the shopkeeper in my example above, learners and teachers choose from among the affordances available to them in that situation, with the overall aim of language learning. We shall see, in Chapter 6 that learners sometimes chose not to use affordances – for example, a transcript of native speaker talk, available to them while performing tasks. Therefore, affordances do not ‘drive’ learning; they are ‘potential’. The use of a particular learning affordance depends on both its availability and on choices made by the individual in the context at that time. The understanding of the process of making such choices and how they relate to task performance is one of the aims of this paper.

In short, from an ecological view, learning is not cognitive, exclusive to a person’s head, but an ever-changing system of relationships between the learner and the language. Van Lier does not deny the cognitive, individual element of learning but insists it is connected to social processes (Van Lier, 2000, p. 258). An ecological perspective borrows the SCT tenet that learning is a process mediated in interaction with others, with the self (private speech), with artefacts (e.g. tasks, textbooks) or with signs. However, the ecological perspective foregrounds the context of the interaction and its available affordances.

2.2.3.2 Ecological research and ecological validity

An early characterisation of ecological validity in research comes from Bronfenbrenner: “An investigation is regarded as ecologically valid if it is carried out in a natural setting and involves objects and activities from everyday life” (Bronfenbrenner, 1979, p. 28). To use more recent terminology, Bronfenbrenner defines ecological research as the investigation of behaviour in a pre-existing, natural context and its pre-existing affordances. He argues that in ecologically valid research “the environment experienced by the subjects [participants] has the properties it is supposed or assumed to have by the investigator” (Bronfenbrenner, 1979, p. 29). Therefore, to investigate a language classroom in an ecologically valid way, the class should consist of the personnel, activities, materials, scheduling (i.e. timetable commitments) learning goals, institutional

constraints, and social mores that participants in that context perceive as being congruent to that context (i.e. its normal affordances and context).

While conducting ecological research, the investigator could examine, among other things, the affordances chosen (and those not chosen) by the participants, the decisions for choosing them and/or the outcomes of these choices on interaction. A researcher could also adjust the context to a certain degree by making different affordances available at different times. This means that an ecological approach could follow the descriptive-phenomenological approach to research used in SCT or the manipulative approach found in cognitive research while investigating language learning in context (Van Lier, 2004, p. 195).

Van Lier laid out four tenets for ecological research:

1. It is contextualised (i.e. situated). The research should focus on relationships within a defined setting – its ecological context.
2. It has spatial and temporal dimensions. Ideally, it would follow a complete, natural life cycle of the focal unit.
3. It can take an interventionist and critical perspective with a view to instigate change. However, this need not be the case for all ecological research.
4. It is ecologically and phenomenologically valid, particularly in terms of an agreement between researchers' and participants' situation definition.

(Van Lier, 2004, p. 193)

These four tenets clarify Bronfenbrenner's earlier definition and add the potential for critical approaches. He, too, argues for the agreement between researcher and participant on the situational definition.

To sum up, researchers engaged in ecologically valid research investigate phenomena (language learning) in the context in which it occurs (a classroom) over a contextually meaningful time span (a task, a lesson, a semester) and in such a way that the people involved feel they are doing the activity under investigation (they feel they are learning English for their sake and not the researcher's).

It would appear that much classroom research could be considered ecologically valid, as

it occurs in the setting where the study takes place. But as Butler (2011) pointed out, much research on TBLT in Asia has been done with tasks introduced for research purposes, brought in by the researcher and a participating-teacher, with only a limited relation to pre-existing context. This study, therefore, sets out to remedy this situation, by investigating pre-existing classrooms and classroom tasks. All activities investigated were chosen for the based on the needs of the students and how the activities fit into institutional curriculum demands. They were not selected for their fit to a theoretical research agenda. In this respect, this study provides an approach for research that will be more meaningful for classroom practice.

2.2.4 This study's perspective

Although treated separately in this review, the ecological approach shares many fundamental ideas with SCT theory, particularly a focus on the dynamics of interaction and the centrality of talk to learning. From these perspectives, task performance itself is learning, as learners jointly co-construct their meanings in the context of the performance. In contrast, from the cognitive, information-processing perspective, researchers demonstrate learning through pre-tests, post-tests and delayed post-tests to show that the anything learnt is now independent of context. The task itself, is an opportunity to notice, consolidate, reconstruct or automatise knowledge (Robinson, 2010) but not 'learning' itself, which takes place within the mind of each individual participant (See Ellis, 2012 for an overview).

However, I find that both the SCT and cognitive perspectives, along with the socio-cognitive attempt to join them (Atkinson, 2002, 2010; Batstone, 2010) limit the range of both questions and explanations to fit a theory. The ecological perspective allows for both qualitative and quantitative analysis (Van Lier, 2004, p. 195) albeit from an SCT worldview. As we shall see in Section 2.5, below, Willis' (1996a) framework for tasks had cognitive origins, but ultimately it was situated ecologically in the context of teaching and learning, and therefore would be better served by an ecological analysis. In addition, some stakeholders in the ecology of any classroom, i.e. government leaders, school directors and such stakeholders want to know what 'typically works' for students. They require 'averages' and 'evidence', as are typically found in studies working in the cognitive framework. In contrast, teachers and learners often want to know what works for them as unique individuals – i.e. what SCT methods usually address. Therefore,

working in an ecological setting necessitates at least consideration, if not use, of concepts from both traditions.

Therefore, rather than stick to a strictly cognitive or SCT perspective, this research works from the premise that all theory is, in essence, metaphor. Metaphors from a Lakoff and Johnson (2003) perspective are means of representing abstract domains (e.g. language learning) through the use of concrete domains (e.g. computers, construction, ecological systems) as expressed in language. Each metaphor highlights some aspects of the phenomenon and hides others, thereby providing only a partial understanding of task performance, while together they provide a more robust picture. So, where any of the theoretical metaphors provides a useful tool or explanation, it will be used.

When taking this, pragmatic approach (Creswell, 2002), there is a risk of mixing metaphors and confusing theoretical perspectives. In order to avoid this, explanations from different perspectives must be considered and addressed in their own right, independent of those from other perspectives. This will be done in this study through first examining quantitative data from a cognitive perspective and then using qualitative data and SCT/ecological concepts to develop the claim.

2.3 Implementing TBLT

This section starts by outlining approaches to bringing tasks into classrooms, then discusses and evaluates the research on how tasks have been implemented. It then moves on to examine how task performances have been researched. It continues by moving through three of the main issues for this research: task input, the effects of task repetition, and how strategy training influences task performance. The chapter finishes with a summary of unresolved issues and an explanation for why investigation is needed.

2.3.1 Using tasks in the language classroom

2.3.1.1 Task-based, task-referenced, and task-supported approaches

Samuda and Bygate (2008) outline three approaches to using tasks across a course of study. In the first, task-based learning and teaching (TBLT) (the ‘strong form’) the

syllabus is defined in terms of tasks and language items are included based on the needs of the learners in order for them to carry out a task. Tasks are selected based on their relationship to real world activities and assessment in the class is based on task performance (p.58). The SSARC (Robinson, 2010) is one way of arranging tasks in this strong form. Willis' (1996a) framework, also advocates this strong form.

The second, Task-referenced learning and teaching (TRLT), is a moderate way of using tasks to teach. Like TBLT, the curriculum is defined by a set of tasks, typically based on learners' needs. Tasks are also used to assess learning, but teachers are not limited to using only tasks in their preparation of learners for those assessments.

The third and 'weakest' form described by Samuda and Bygate (2008) is: Task-supported learning and teaching (TSLT). In this form, tasks are options available for teachers to use to enhance the learning experience, but the curriculum and assessments may well be based on other criteria.

This distinction highlights one of the problems in the literature, which shall come up again, below. This is the issue of different researchers applying the same label, typically TBLT, to approaches to teaching that are not task-based. They may even apply the label 'task' to things that are less 'task-like' (J. Willis & Willis, 2007). See, for example, Ellis' (2009a) response to Seedhouse (2005), Swan (2005), and Sheen (2003) for discussion of this issue.

2.3.1.2 Focused and unfocused tasks

Ellis (2003) offers a dichotomy between focused and unfocused tasks. An unfocused task is not designed to draw attention to a specific language form. A focused task, in contrast, attempts to do just this, and it is typically focused on a grammatical structure. This engagement can be productive or receptive or both. The focused task must still conform to the definition of task, so "the targeted feature cannot be specified in the rubric of the task" (p. 16).

Focused tasks, therefore may be more helpful for teachers and syllabus designers, as the targeted features may be pre-identified for their pedagogical or theoretical relevance. Tasks can be given focus by making use of certain items 'task essential' (Loschky & Bley-Vroman, 1993). This means arranging the task so that certain forms must be used in order for the task to be done successfully. In practice, there are a limited number of items

which can be targeted in this way as learners can avoid using the targeted forms by using simpler forms, telegraphic speech or even gestures. Another way to focus a task is to make the task about language. Consciousness raising activities (Fotos, 1994; D. Willis & Willis, 1996) and dictogloss (Wajnryb, 1988) are two common ways of doing this.

2.4 Task implementation and task performance

This section will focus on task research that deals with the way tasks are implemented in classrooms and with ways of operationalizing and characterising task performance. The focus on classrooms is needed because the present study is situated in a classroom. That being said, there are not actually that many ecologically valid (Van Lier, 2004) classroom studies. Many studies that were done in classrooms were the result of teachers allowing researchers access for a few weeks and the participating students doing tasks that were not originally intended to meet their pedagogical needs, but which were designed to address research questions the investigators wished to answer (Butler, 2011). This does not mean that research conducted under experimental conditions or these ‘parachuted’ tasks are not valid and important, or that their results would be different if they were done in classrooms (Gass, Mackey, & Ross-Feldman, 2011). When such research directly impacts the present study they shall be reviewed, but the preference is for situated studies whenever possible. I start by looking at measures of task performance. Typically, these are dependent variables such as measures of the complexity, accuracy and fluency (CAF) of learners’ output, or counts of language related episodes (LREs).

2.4.1 Observing and interviewing: Teacher perceptions and behaviours

Teachers are the ones who bring tasks into the classroom, or the ones who implement those from prescribed texts, so how they do this and how they feel about tasks is a reasonable way to discuss task performance.

Generally, teachers in school settings have had difficulty implementing tasks. Carless (2004) observed and interviewed three Hong Kong primary school teachers as they tried to implement a new governmental directive calling for TBLT. He observed them across six consecutive lessons each and audio-recorded these lessons. The participants were

observed seventeen times each over the course of the study and were interviewed six times. He found that they had mixed feelings about L1 use, that some was helpful but there was often inappropriate use, for example, arguing. He found that teachers had difficulty balancing tasks with their feelings of good classroom discipline and noise level. Finally, he found that the tasks were often dominated by one or two speakers and for some students, and in some tasks, there was little L2 production. He observed that the teachers had “filtered” the tasks to something more like the production phase of a present – practice – produce (PPP) lesson because of challenges they had in implementing tasks (p. 658). For example, one critical episode was an activity designed to specifically use forms presented immediately before doing the task. It was then practised with the whole class and then the students were asked to do the ‘task’ (Carless, 2004, pp. 650–1). This does not meet the definitions of ‘task’ (Sect. 2.1.1) as its purpose was to use the picture to elicit a practiced grammar form. Carless offered no insight as to why the teacher chose to do this activity in this way. Similarly, a second critical episode, (pp.651-2) had the structure of an information gap, but because the students were drawing pictures of the playground, something they had all seen before, there was no ‘gap’, therefore no new information was needed by the student who was not looking out the window, so there was no ‘challenge’.

Carless’ study highlighted some of the main problems teachers face when implementing tasks in classrooms. It also raised issues at the interface of tasks in classrooms and task based research. Carless mentioned but did not address the issue that teachers had changed the meaning of task in this study, and it is possible that his findings, particularly with respect to student-behaviour really reflect a present – practice – produce (PPP) approach to teaching and not TBLT.

Carless also used interviews and observations to address the suitability of TBLT vs. PPP in Hong Kong in follow-up studies (Carless, 2007, 2009). From this data he found that teachers saw grammar exercises as boring for students; and, although students liked tasks, these may not provide learning outcomes that teachers want (Carless, 2009, p. 55).

Carless (Carless, 2004, 2007, 2009) reported that teachers adapted the tasks to suit their beliefs about language and learning and, particularly when in charge of higher grades, the institutional demands for testing. Littlewood (2007) demonstrated that Carless’ (2004) report was common in the wider Asian context.

This observational and interview-based research used teachers' perceptions of tasks as the basis for measuring task performance. When considering tasks' usefulness, the teachers considered the amount of L1, student behaviour /discipline, the amount of L2 – often a targeted and practiced structure, students' feelings, and the learning potential of the task.

2.4.2 Experimental research: Complexity, Accuracy and Fluency (CAF)

A large body of experimental research has focused on the relationship between aspects of task design and the complexity, accuracy and fluency (CAF) of learners' language production. Participants are given a task under certain conditions and their performance (usually their talk) is recorded and analysed along a number of dimensions. Ellis and Barkhuizen (2005, p. 139) say that accurate language is language that follows the rule system and represents the speaker being conservative and using known items. Complex language is at the edge of a learner's interlanguage and represents more of a communicative risk – almost opposite to accurate language. Fluent language is free of hesitations and false starts, and produced at a reasonable speed.

Research on CAF formed the backbone of the TCF and LAC models of tasks described above (Robinson, 2001; Skehan, 2001). As a result of the design or implementation of a task, the CAF of the input was predicted to behave in certain ways. A large amount of research using CAF measures has been conducted specifically to test the effects of different conditions with respect to one or both of the models (For example: Foster & Tavakoli, 2009; Housen & Kuiken, 2009; Jackson & Suethanapornkul, 2013; Robinson & Gilabert, 2007; Robinson, 2001, 2011b; Shiau & Adams, 2011; Skehan, 2009; Tavakoli & Foster, 2008). The results have been inconsistent across these studies leaving Skehan (2009) and Robinson (2011a) both calling for more sensitive measures. As Ellis and Barkhuizen (2005) also pointed out, there have been a large number of ways of measuring CAF, none universally agreed upon, and this has contributed to the problem. They also say that a resolution between the two theories may be impossible due to the presence of too many other variables (Ellis & Barkhuizen, 2005, pp. 144–5).

2.4.3 Language Related Episodes (LRE)

The construct of language related episode (LRE) originated in socio-cultural theory as an extension and development of the cognitive measure “Negotiation of Meaning” (Long, 1996). Swain and Lapkin (Swain & Lapkin, 1998, p. 326) define an LRE 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.” This talk about their language represents an instance of noticing in interaction.

LREs are a type of interactional feedback, either with the self (self-correction) or with others. They have been shown to be influenced by factors such as planning time (Philp, Oliver, & Mackey, 2006), task type (Swain & Lapkin, 2001), proficiency (Leeser, 2004) pair dynamics (Storch & Aldosari, 2012; Storch, 2002) and group dynamics (Philp et al., 2010).

LREs in many studies are separated into types. Divisions along the focus of the LRE, grammatical and lexical (vocabulary) and whether or not they are successfully resolved being the most common (e.g. Kim & McDonough, 2011; Kim, 2008, 2013; Swain & Lapkin, 1998, 2001). LREs have been taken to represent learning opportunities and therefore coded and counted as measures of the learning potential of a task, however, as Fortune and Thorp (2001) argue, this coding and counting is simplistic, as LREs can occur over many turns, have other LREs embedded within them, or be started at one point and finally resolved much later. Storch (2008) found that LREs that were deliberated over, or elaborated, resulted in better learning than ones that were engaged with in a limited way, suggesting that not all LREs are equal.

The LRE represents a point in the interaction in which learners address the language they are using, that is, it is a point when attention is directed to language. If we follow Schmidt’s (1990) claim that learning cannot happen without noticing, then LREs are evidence of the number of learning opportunities (Crabbe, 2003, 2007) in a particular interaction and task performance will then reflect the learning potential of that task for those learners.

2.4.4 L1 use in tasks

One of the more controversial issues in task based learning and teaching (TBLT) and in foreign language teaching in general is the use of the students’ first language (Alegría de la Colina & García Mayo, 2009; Carless, 2004, 2008; Hall & Cook, 2012; S. Park & Manning, 2012; Pinter, 2005; Storch & Aldosari, 2010; Swain & Lapkin, 2000). The current consensus is that L1 use is a learner choice and is not necessarily a problem. Disagreement may arise about the amount, timing and function of the L1 use. In their

more recent framework for TBLT, the Willis' (J. Willis & Willis, 2007) allow for L1 use among lower level learners during the task and planning stages but recommend learners use only the target language, L2, in the public report stage. Similarly, Hall and Cook (2012) discuss the issue and place of L1 use not as an all-or-none, but as a continuum of practice in which different needs are met using the different languages.

Pinter (2005, 2007), who was working with young children whose English proficiency was low, had her students do tasks first in L1. She then had them repeat the task, this time in L2. This approach had the benefit of letting children know what was expected of them in terms of outcome without the added pressure of using English. She stated, "Asking the children to carry out the tasks first in their L1 seems a good way to judge whether the tasks are likely to work in L2" (Pinter, 2005, p. 123). Her approach was unique in this respect; the entire task was done first in L1. It is questionable if such an approach would be useful for older or slightly more proficient learners, but L1 met the needs of Pinter's learners, and allowed for superior L2 production in the repeated task.

Rather than focus on using or not using L1, some research looked at L1 as an emergent phenomenon in task performance. That is, this research assumed L1 might be used and sought to find out the circumstances in which it was used. One finding specific to TBLT was that L1 use is task-specific, that is, some tasks seem to promote or require more L1 than others (Alegría de la Colina & García Mayo, 2009; Carless, 2008; Storch & Aldosari, 2010).

Alegría de la Colina and García Mayo (2009) worked from the assumption that using L1 would allow students to work at a higher level than if they did not. They looked at three tasks: jigsaw, text reconstruction and dictogloss. They found that L1 use happened in different amounts and at different points in the task. For example, the jigsaw task used the least L1 overall, but in this task once learners finished talking about the oral version they then had to jointly compose a written version. At this point, they had the most L1 metatalk, as they moved from oral L1 to written L2 (p.340). Meaning was made in L1 and transposed to L2. The other finding from this study was that there was almost no off-task behaviour, so L1 use was not an indication of non-engagement with the task as feared by the teachers in Carless' (2004, 2008) studies.

Storch and Aldosari (2010) addressed the complaint of teachers that students will use the L1. They investigated the use of L1 across three tasks by three different proficiency-level

pairings: High-high, high-low and low-low proficiency. They reported that participants used rather little L1. And that of the three tasks they investigated: jigsaw, composition and text editing, the most L1 was used in the text editing task. They also found that task type was more influential on the amount of L1 than proficiency level pairing, but there were individual differences across pairs. Some pairs used less than 5 per cent of their turns in L1, and task type had no effect on these 'low L1' pairs. Moderate and high L1 pairs were impacted. Storch and Aldosori (2010) report qualitative differences across pairs of learners in terms of type and function of L1, indicating that L1 use is more idiosyncratic, with reasons that lie within the individuals and their relation to the partner, rather than being determined by the task, or the proficiency level of the partner.

To summarise the research on L1 use and tasks, there is little push among current research to force L2 only policies. L1 has been found to help the language classroom, and some tasks seem to predispose more L1 than others. Carless (2008) reported a teacher-trainer saying, "A thing that is crucial is the wonderful paradox that if the task is a good task, students do it in Cantonese" (p. 355) So, if a task is engaging and interesting, students will not try to learn from it, but enjoy it. The difference between Carless' interview study and Storch and Aldosori (2012) and Alegría de la Colina and García Mayo (2009) might lie in the fact that the latter two deal with university students, people who are already successful students and may be more inclined to learn the L2, know how to learn it, and what is expected of them, not to mention more physically and emotionally mature than elementary school children who Carless' informants were teaching.

The current study will be investigating university students, so it is expected that there would be similar amounts of L1 to Storch and Aldosori (2010), and that L1 use will be dependent on both task and individual factors. However, neither Alegría de la Colina and García Mayo (2009) nor Storch and Aldosori (2010) took steps via strategy training or implicit awareness raising to influence the amount of L1. In addition, neither of these two studies on L1 use report off-topic L1 at all, whereas Carless' (2008) informants complain about it a great deal (p.333-4). In the two studies reported here that were done with university students there may have been no off topic L1, as the research was carried out by volunteers, but in a classroom, there will be students who do not participate fully. The extent to which this disrupts their task performance has not yet been established or if implicit or explicit awareness training (See Section 2.8) could reduce it.

2.5 Willis' (1996) pedagogical framework for TBLT

The Willis framework for TBLT starts from the basic premise of 'meaning before form', but it does not ignore the need for improving formal accuracy. The model incorporates the main task and other activities into a task cycle which includes other activities, both in the pre-task before the main task and in the post-task after the main task. One aspect of the 'meaning before form' approach is that language form is addressed after the task.

The pre-task phase introduces the topic and vocabulary and may involve the use of a model of the task; the task cycle involves three parts – the task, also called, 'activity', planning, and reporting to the class; a post-task phase that includes language analysis; and finally controlled practice. The framework is in Figure 2.1.

2.5.1 The task cycle

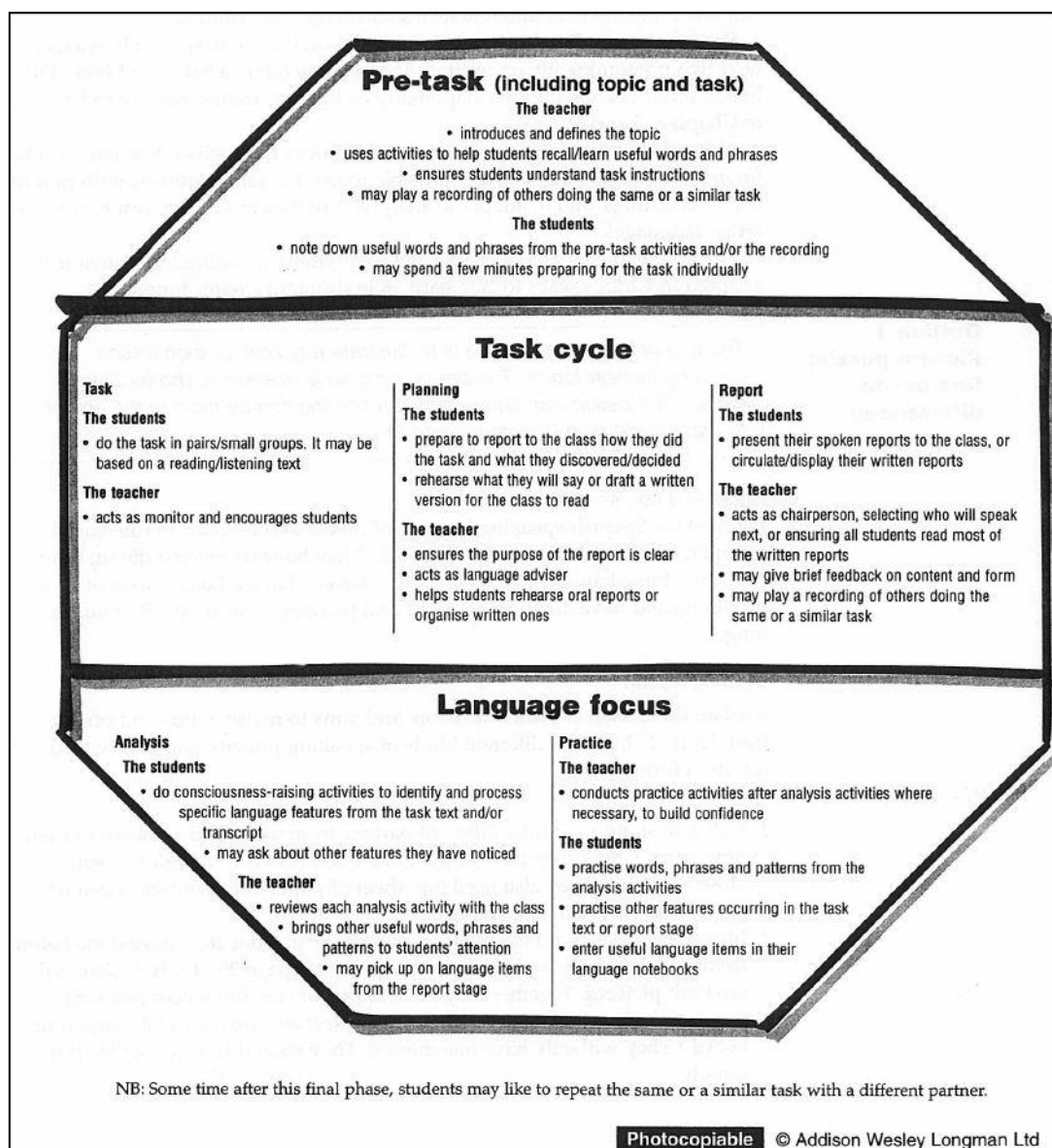
As can be seen in Figure 2.1, the task cycle calls for different behaviours and different roles for the instructor and learners at all stages. However, there has been no research on whether or not this happens. More importantly, the second two parts of the cycle: the planning and report stage, are emphasized in the Willis model as driving a push to more accurate language, but they have also not been researched. Planning has been operationalized in the literature as 'pre-task' or 'online' planning (Ellis, 2005, 2009b) and not as a move that involves making the learners' informal speech (i.e. within the group) ready for public listening. Likewise, the effect of post-task conditions on task performance has not been widely studied.

2.5.1.1 The task stage

In this part of the lesson, the learners do the task and work toward an outcome. The teacher has the role of monitor in this stage and is not supposed to 'teach'. Willis implies that this 'teaching' may actually be interfering and suggests that teachers do not give too much assistance. Willis suggests that during the task, teachers should be making sure all students are doing the right task (an allusion to the task-as-workplan vs. task-as-process distinction discussed above), encouraging all students, forgiving of grammatical errors, helping only if there is a major communication breakdown, noticing the individual performances within the groups (who is dominating/hiding), noticing switches to mother-

tongue and reasons for the switch, and time-keeping (Willis, 1996a, p.54). The role of the teacher is a facilitator.

Figure 2.1: Overview of the TBL Framework (J. Willis, 1996a, p. 155)



Willis claims that the task itself is not enough to promote language learning. This appears to be in direct contradiction to a great deal of research, for example, Robinson's cognitive hypothesis (Robinson & Gilabert, 2007; Robinson, 2007, 2011b). However, she claims: "The task component, as we have seen, helps students to develop fluency in the target language and strategies for communication" (J. Willis, 1996a, p.54). But if there is no public presentation, a change in what Robinson calls, "task conditions" (Robinson, 2007, 2011b), then, according to the Willis model, language learning will suffer. She claims that if there is no public performance of the task: (1) Some learners will use L1 and not try to

develop their L2, especially if their L2 is weaker than their group members; (2) Some learners will develop communication strategies and not L2; (3) Some learners will overly focus on finding/using the right words and not fit the words into the discourse; and, (4) there will be more concern for lexis than grammar (i.e. there will not be any LREs with respect to grammar) (J. Willis, 1996b, p. 55). However, research to prove the effectiveness of the public report on improving student performance is limited (But see Ngyuen, 2013).

With respect to the use of L1, there seems to be some confusion about its role in TBLT. Willis admits that L1 can be used in the task, but not during the report (Willis & Willis, 2007). Allowing L1 in the task and planning phases of the task cycle may well help participants express their ideas, but it will impact on the number and kind of learning opportunities afforded by the *process* of doing the task.

2.5.1.2 Planning and reporting stages

Willis (1996a) claims that in order to promote more accurate language use, teachers need to adjust the task cycle to create a need for accurate language. She says, “The report stage is when groups report briefly in spoken or written form to the whole class on some aspect of their task, such as who won the game, how their group solved the problem, or two or three things they found out from each other. In doing this, students find themselves in a situation where they will be talking or writing for a more public audience” (p.55). She says that because “Language used in public is normally planned, final and permanent” students will want to use their best language, avoid mistakes and make sure what they want to say will be comprehended by their classmates (p.55).

The planning stage, immediately before the report, is the point of the task cycle where learning is pushed according to Willis.

In planning their report, they [students] have time to create anew, experiment with language and compose with the support of their group, teacher, dictionaries and grammar books. *And it is this process that is likely to drive their language development forward and give them new insights into language use* (J. Willis, 1996b, p. 56).

What has not yet been offered is an ecologically valid account of learners’ in-task talk showing both the change in language to a more accurate version and awareness by the

students of the need to do so. This framework also allows for a written or spoken report, yet whether one leads to different learning outcomes than the other has also not been researched.

Similarly, her claims about the report stage remain unproved:

The report stage gives students a natural stimulus to upgrade and improve their language. It encourages them to think about form as well as meaning; accuracy as well as fluency and to use their prestige version of the target language (J. Willis, 1996b, p. 64).

Note that the report phase is not necessarily asking the learners to repeat the task in public for the whole class to watch. It could be, but it could also be something related to the task, such as to discuss an aspect of it or to summarize it. In this framework, the report is not a task repetition but a form of self-analysis, or even peer review. In essence, the report phase could be used as a follow-up task and not a repetition or rehearsal. Of the report, Willis (J. Willis, 1996b) says, “In itself it [the report stage] 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). Here she clearly relates language learning to factors external to the task itself, something acknowledged, but not considered in task design, by some task researchers (E.g. task conditions in Robinson, 2001, 2007, 2011b) and highlighted by others (E.g. alignment in Atkinson et al., 2007; and language socialization in Duff & Kobayashi, 2010).

By separating the task and the public report, the framework highlights the distinctions between product (the final report) and process (the task and planning) expressed in Samuda and Bygate’s (2008) definition of task. The potential lack of alignment of the task with the report implies there may really be two tasks going on in one cycle – doing the task, then preparing to talk about it publicly. This potential lack of alignment is an ecological factor and may well affect task performance.

2.5.1.3 The pre- and post-task phases

Keeping in mind that the Willis’ framework for implementing tasks in classrooms is based on the premise of “meaning before form”, the pre-task is an activity or set of activities that provides the learners with the language they need to be able to express meanings. In general, this means vocabulary development of some sort. It may also

involve providing a model for learners to follow. Providing words and meanings that participants will need to use while doing the task makes intrinsic sense, according to Willis, and the benefits of pre-task modelling have been documented (Kim & McDonough, 2011), we return to the issue of pre-task input in section 2.2.

The post-task phase was identified by Willis as being the best point in the framework to have students work on the formal accuracy of their language (J. Willis & Willis, 2007, pp. 113–134; J. Willis, 1996b, pp. 101–115). The activities in this stage are typically in the form of consciousness raising (CR) activities (Rutherford, 1987; D. Willis & Willis, 1996). Although CR activities have been shown to aid in language learning, the value of placing them after the task has not yet been empirically substantiated. This placement of instruction and activities related to form after the task was based on Schmidt's (1990) idea of learners needing to notice a gap between what they are capable of saying in the target language and what they wish to say. Having noticed issues with their language when doing the task, learners would be primed to learn from instruction or input that would help them fill in the gaps if this input and instruction came after the task (J. Willis & Willis, 2007; J. Willis, 1996b).

2.5.2 The Willis' framework and syllabus design

Willis' framework has arisen from pedagogical practice. As such it is popular among teachers (see Edwards & Willis, 2005). It divides a lesson into stages based on the three task phases across which students will be pushed to produce more accurate language if they are required to present a report publicly; and it requires instructors to draw their learners' attention to the formal aspects of language.

Willis' framework shows similarities to other pedagogical models, for example the 'Four strands' notion of Paul Nation (Nation, 2007) in which meaningful input, meaningful output, deliberate study and fluency development (i.e. increasing time pressure on output) are considered the essential parts of a language learning course. In Willis' framework, the pre-task and task provide input and output oriented to meaning, the post-task provides deliberate study. Fluency development fits into Willis' framework if the task is repeated, although Willis considers this optional (See Figure 2.1, above).

Willis (1996a, 1996b) contrasted her framework with the Present-Practice-Produce (PPP) approach. In PPP, the teacher presents and explains language points first, then has

students practise the target items through drills or exercises before they do more open, production, activities which require the use of the target items. PPP is a ‘form before meaning’ approach, and in Willis’ view, was not as effective for learners as the task cycle (J. Willis, 1996a, 1996b)

The framework, however, has received little attention from researchers in terms of the predictions made at the various stages. For example: What type of report pushes planning talk to focus on form? Will language learning be improved if the task and the report match each other closely? Does doing the post-task CR activities drive learning better than doing the CR in the pre-task phase? Does a pre-task model help task performance? Under what conditions will repetition help task performance?

One possible reason for this lack of research is that the framework has not clearly proposed criteria for developing and sequencing tasks. Willis suggests using Skehan’s (1998) LAC model as a checklist to refine teacher intuition about a task and then look at the complexity of task-related texts as a guide for determining the complexity, and therefore the sequence, of tasks (J. Willis & Willis, 2007, p. 186). Willis suggests incorporating the Common European Framework (Council of Europe, 2001) to create a task-based syllabus but gives no further guidance. Her framework relies on teachers, in their setting, to know through experience what would be an appropriate task and level of difficulty for their students’ learning needs. This situating of tasks in the contexts where they are to be subsequently used suggests that the framework actually has an ecological basis. So, despite Willis’ attempts to situate her framework in cognitive theory, particularly her emphasis on pushed output driving noticing in order to increase learning, her framework, it is, at heart, an ecological framework – cognition is situated in a context. This, therefore, suggests that any approach to researching it should follow an ecologically valid approach (Van Lier, 2004). However, researchers working from the Interaction Hypothesis approach have largely ignored ecological settings in favour of experimental designs. They favour the precise control of cognitive variables believed to push complexity, accuracy and fluency in an attempt to develop a theoretical basis on which to pre-specify a syllabus without referring to the learners or teachers who would use it (e.g. Robinson, 2001; Skehan, 1998). This too, has contributed to the lack of interest in Willis’ framework.

2.6 Input and task performance

From cognitive perspectives on task based learning and teaching (TBLT), input has a variety of roles in tasks including: (1) input tasks which rely entirely on the learner processing input, with no pressure to produce language (e.g. Shintani, 2011); (2) pre-task modelling of the task (e.g. Kim & McDonough, 2011); (3) something to be understood and/or manipulated in the task (e.g. Seedhouse & Almutairi, 2009), including task instructions and prompts (Ellis, 2003; Newton, 2013); (4) a source of post-task information to be examined and to which students can compare their performance (J. Willis, 1996b) although the effects of post-task input have rarely been studied; and finally, (5) output from an interlocutor is input for the hearer (Long, 1996).

In this last respect, input has been most extensively studied as a source of learning. The interaction hypothesis claims that negotiating for meaning in task interaction will lead to modified input that has been brought to a level understandable to the interlocutors, and therefore able to be learned. This input is also the source of and means of resolving LREs. It is in this sense that input has been studied in greatest detail (See Gass & Mackey, 2007; Mackey, 2012 for overviews). This study is interested in the role of input as a pre- and post-task resource for learning.

2.6.1 Input in the pre-task

Using input before the main task allows the topic to be set up and students' background knowledge and opinions to be activated and gives an opportunity to introduce vocabulary about the topic (J. Willis & Willis, 2007, p. 21). They suggest that for some tasks, presenting learners with a native speaker model before they do the task may be helpful for them. As Willis argues:

By listening first to recordings of one or two pairs doing a similar task students will get a clearer idea of what they are expected to do, and how. They may get ideas about what to say themselves, and of the general shape of the interaction. It may also help them recall useful words and phrases. (J. Willis, 1996b, p. 90).

Here, Willis lays out the general function of pre-task input: to show the task procedure, to give ideas with respect to content, and to stimulate memory of previously learned words,

along with the introduction of a small amount of new vocabulary in this phase.

The first function of the pre-task, to give a clearer idea of task procedures was investigated by Kim and McDonough (2011). They studied the effect of having the instructor and the researcher model a task on student output. The model performances in the study were written in advance in order to ensure they contained examples of LREs targeting both vocabulary and grammar that were correctly resolved; and, they contained examples of collaborative interaction including peer feedback, questions and sharing of ideas. This is not the same as Willis' suggestion of native speaker models doing the task because it controlled the content of the recording and provided examples of cooperative learning behaviour, and lost the false starts, hesitations and other imperfections of native speaker speech that Willis recommended learners hear because they need exposure to authentic language (J. Willis, 1996b, p. 89). In essence, Kim and McDonough (2011) modelled an idealised task performance and in doing so raised the learners' awareness both of what was expected in terms of task outcomes, and what was expected in terms of task process. They found an increase in the number of LREs and more collaborative pair dynamics in the classes that received the pretask models. They also found that different task types were affected differently by modelling. This suggests that pre-task models have a facilitative effect on the use of tasks as learning opportunities by students. They did not investigate the extent to which the language in the models was used by their students as they performed the tasks. They mentioned, as an anecdote, that classes that received models mined (cf. Boston, 2008) the materials more generatively, whereas those that did not receive a model, lifted items verbatim. However it appears that Kim and McDonough (2011) were talking about items from the task prompts and not items from the video model.

Kim and McDonough's (2011) version of modelling, in line with earlier research (e.g. Swain & Lapkin, 1998, 2001), found a facilitative effect of pretask models on collaborative pair dynamics and the number of LREs. But as their study was not about language mining, they did not discuss this in detail.

Boston (2008) investigated the mining of pre-task input by Japanese false-beginners in task performance. He investigated the question of whether listening to an audio recording before doing the task would enable learners to mine items from the recording in their task performance and would this be comparable to visually presented input. He found that his

learners mined items from the visual input, essentially, the task instructions written on the paper in front of them, and they did not mine the content of the recordings, even though they had listened to the recordings multiple times. He did not give his learners the scripts of the recordings to read or examine, however, so the students might have forgotten the form of the items in the recording. Even though his students had heard the recording several times, because the recordings were processed for overall meaning and not for specific vocabulary, expressions or structure, the students had no need to remember the formal details of the items in the recording. That Boston's students mined items written in the task instructions, suggests that reading input may enhance the possibility of these items being mined.

These two studies exemplify how input has been used in task-based teaching: As a guide for task behaviour and only minimally as a language resource. Kim and McDonough (2011) only briefly mention language use from the tasks, since their focus was on how models provided a blueprint for the task itself and for issues that were expected to arise within the task. As such, the models were scripted to ensure the interactional features they wished the students to use were present in the models. They were not authentic in the sense of providing exposure to false starts, hesitations and elided speech. Boston's (2008) study looked at the use of pre-task items in the task itself and found that aural input had little impact on the items used in task performance.

Neither study attempted to pre-teach and practice vocabulary or grammatical forms in the pre-task. Boston (2008, p. 72) suggested that teachers could "seed" items into the input with the intention of stimulating students to implicitly mine and use them in their talk, arguing that this was not the same as drilling a form.

This current study will address this gap by providing model performances of the task by native speakers (NS) in oral and written forms. The students will then be given consciousness raising (CR) activities (Ellis, 2003; D. Willis & Willis, 1996) that require them to analyse these models in order to pre-teach task-useful vocabulary that *could*, not *must* be used (Loschky & Bley-Vroman, 1993). Following Boston (2008), some classes will get the input (model and CR) prior to the task and others will get it after. It was expected that the pre-task input groups would show similar results of LREs to Kim and McDonough (2011), that is, more LREs when the task is modelled. It was expected that there would be more items when participants had pre-task input, similar to Boston (2008).

2.6.2 Input in the post-task

Willis (1996) says that for closed tasks, students should listen to input after doing the task themselves and compare their performance to the native speakers because “they may notice words and phrases they could have used themselves” (p. 90). They also suggest doing Language Focus activities to enhance knowledge of form – typically based on consciousness raising (CR), but other methods are allowed in the post-task as well. The reason given is that learners will learn if they have noticed the gap between what they know and what they need in order to express themselves fully (Schmidt, 1990, 1994; J. Willis & Willis, 2007; J. Willis, 1996b). To test if participants will perform a task better when given post-task input than when given pre-task input, the task will be repeated. So, we turn now to a discussion of the literature pertaining to task repetition.

2.7 Task repetition

One of the underrepresented task implementation conditions in the literature is task repetition. Willis (1996a, 1996b) suggests that tasks could be repeated with new partners after the language focus in order to consolidate learning (See Fig. 2.1). A theoretical argument for repetition, as proposed by Ellis (2003), is that the first performance allows the learner to decide on the meanings they wish to convey, and then in their subsequent task performances, they have more attentional resources available to address formal aspects of their performance. In the limited literature on task repetition that exists, there are issues with how task repetition is operationalized and in the predominant use of monologic tasks, not interactive tasks (Kim, 2013).

Early studies on task repetition treated repetition as a form of ‘familiarity’. Plough and Gass (1993) found that pairs of students who were familiar with the task had better interaction, they interrupted less, interacted more via clarification requests and comprehension checks.

Lynch and Maclean (Lynch & Maclean, 2000, 2001) developed and reported on a poster carousel activity for use with medical students learning English. Students had to work in pairs to prepare a poster about a medical journal article. Then one member of the pair

stayed with the poster and others moved around the room. The person who remained had to answer questions from a learner who moved from group to group, a 'visitor'. The process was repeated until all participants had visited each poster (6 repetitions) and then the roles were switched. They found that the weakest student's accuracy improved in syntax, lexis and phonology, but she did not report improving. They also found that their strongest student reported planning changes to her language in advance. In addition to this she made unplanned changes. When Lynch and Maclean (2000) made a detailed examination of this stronger student's talk, they found that her explanation of a new concept became more detailed and then trimmed of excess wordings across the repetitions. They view the repetition as more like a rehearsal or a retrieval, and not a repetition because the task changed due to the visitor's questions (p.227).

Mackey, Kaganas and Oliver (2007) investigated task familiarity and child task performance while simultaneously studying planning time. They operationalized familiarity by training the children how to do the tasks before the study. Half of the children were trained to do one task, and the others were trained on the other. They followed a similar strategy with task content. So, although each child did four tasks, two of one type and two of the other, and each child used both familiar and unfamiliar topics, Mackey et al (2007) did not directly investigate repetition. But, they found that when tasks had unfamiliar content and unfamiliar procedures, the children engaged in more confirmation checks and clarification requests and gave each other more corrective feedback. If the task procedure was familiar, but not the topic, their learners had more opportunity to use the feedback, and if the topic and the task procedure were both familiar, they used the feedback more. From this study's perspective, task repetition could be seen as a kind of familiarisation procedure, and task repetition should have a similar impact as familiarity training.

Shintani (Shintani & Ellis, 2011; Shintani, 2012) reported the effect of task repetition on young learners. She was investigating two teaching methods, production based and comprehension based learning. She operationalized comprehension based learning as using input tasks (Ellis, 2012), tasks in which learners were required to understand input and perform actions but are not required to produce language, but may do so voluntarily if they wish. Of interest here is that she repeated the same three input-tasks across six lessons with a group of very young learners. She reported that the children enjoyed the tasks, increased their voluntary output, and found evidence of incidental learning of the

plural ‘-s’, but only in students who had never studied English prior to joining the class. The possibility of input task repetition as a medium for incidental vocabulary and grammar learning while simultaneously aiding output production is intriguing and needs to be investigated for older, more proficient learners, to see if it was the input or the voluntary output that pushed the learning.

Kim (2013), investigated two types of repetition: task repetition and procedural repetition. Task repetition referred to the exact, or nearly exact, re-doing of the same task on the same topic as exemplified in Shintani’s studies (Shintani & Ellis, 2011; Shintani, 2012). Procedural repetition referred to doing tasks of the same structure but with different topics, and therefore different lexical content, similar to Lynch and Maclean’s idea of retrieval (Lynch & Maclean, 2000, p. 227). She examined if task and procedural repetition would have different impacts on learners’ attention to linguistic forms, and how the learners and teacher would react to the different types of repetition.

She had pairs of students do the same information gap task three times (task repetition) or use three different information gap tasks once each (procedural repetition). The tasks were done over three lessons in a one-week period. She also surveyed the participating students and interviewed the teacher. Her results showed that the task repetition group increased the number of grammar-related LREs (LREg’s) they produced over the three repetitions, indicating a move to more accurate speech. The procedural repetition group used more LREs overall. This group also required more teacher intervention to resolve their LREs, whereas the task repetition group tended to resolve the LREs themselves, again pointing to a benefit of task repetition.

Although the procedural repetition group liked the greater variety of topics, the teacher reported some reluctance from her students to do the task the third time in the task repetition condition. The teacher also reported a desire for sample vocabulary and sample sentences to be included with the tasks to help the learners, that is, some task-related pre-task input (See section 2.4).

One confounding issue in Kim’s study was that the pairs were changed each time the tasks were done. Interlocutor familiarity has been shown to influence task performance and changing the interlocutors may well have influenced the results (Plough & Gass, 1993). This is an issue that will also arise in the current study (See Ch. 4).

In another classroom-based study, White (2011) investigated the effect of immediate procedural repetition of the dictogloss task. He found that there was more noticing and use of targeted lexical items in the repeated task. White did not counterbalance his design so there may have been a topic effect, but the result is intriguing because he had targeted lexical items in the dictogloss that had been introduced a week previously as part of the course, and found that learners picked them up better with the repetition. The implication for classroom practice, particularly if he had counterbalanced the design, would be to repeat, not one lesson later as in Kim (2013), Shintani (2011) and Mackey et al (2007), but within the same lesson, as in Lynch and Maclean (2000, 2001).

One thing not considered in the research is the number of repetitions. White (2011) had two, Kim (2013) three, Mackey et al (2007) had two in each of two conditions, Shintani (2011) six and Lynch and Maclean (2000, 2001) also had six. At what point do repetitions become ineffective or boring to students (Kim, 2013), is a question that also needs addressing.

Another pertinent question is what happened between repetitions. In another classroom based study, Hawkes (2011) added a form focus activity between two repetitions of a task. He investigated three different tasks: opinion gap, describe and draw, and a timed discussion. Participants were given a model of advanced speakers doing the same task in the pre-task phase. They then did the task, and then did a consciousness raising activity similar to ones described by Willis (Hawkes, 2011, p. 330). He then had the participants do the same task with the same interlocutor, without looking at their notes. He examined the number of corrections per AS unit (Foster, Tonkyn, & Wigglesworth, 2000) and found that the repeated performances all had higher numbers of corrections. He also noted that the corrections were predominantly aimed at the items targeted in the focus on form. Hawkes concluded that repeating the task after an intervening focus on form would be a way to enhance the grammar teaching and learning potential of tasks.

To sum up this section, task repetition has not received much attention in the literature. It has, however, been shown to facilitate focus to formal accuracy in a number of studies across a variety of tasks. Some questions remain unanswered. One of which is the longer-term impact of task repetition. Only Shintani (2011) examined multiple repetitions across multiple lessons and both a post-test and a delayed post-test. Her participants were young children, and so this raises the questions of whether older learners get bored by doing

multiple versions of a task as Kim's (2013) participating teacher reported, and the extent to which older learners would benefit from doing multiple versions of a task across several lessons.

Many teachers do not have time to repeat a task twice in one lesson. Would Hawkes' (2011) idea, following Willis (1996b), of having a focus on form between repetitions still hold if the second repetition were delayed until at the start of the next lesson?

Finally, if learners behave as reported by Kim (2013), that is, they got bored and said "again?" when the same task was used (Kim, 2013, p. 17), what approaches can teachers use to re-engage students with the task so they get maximum learning benefit from doing it?

2.8 Learner awareness and task-based learning

Ellis (2003, p. 8) remarked that task designers needed to use "sleight of hand" when making tasks, because students need to focus on the outcome of the task in order to stimulate communication, but the real purpose of the task is not to finish the task, but to engage language use in such a way that language learning is promoted. There is a worry among task designers, particularly for focused tasks, that if the purpose of the task is announced, students will simply display knowledge and not communicate. The opposite worry is also present; students will over-focus on finishing the task and use L1, gestures, telegraphic speech (minimal L2) or a combination of these, in order to finish the task, thereby not learning from their interaction (Carless, 2004). With focused tasks, there is also the possibility of learners using language other than the targeted form to finish the task (R. Sato, 2010). This raises the question, that if students are doing a task, do they realise what the purpose of the task is? And, if we let students know the interactive purpose of the task, do they interact differently, more specifically, do they interact in a manner that more closely matches the language learning goals of the task?

To address this issue, I will review studies that provided learners with training on aspects of interactive learning.

Working from a sociocultural perspective, Ewald (2004) investigated students' and

teachers' perceptions about using small group discussions for L2 Spanish learning. She gave students a 50-60 minute forum that included journals, "mostly agree/mostly disagree" questionnaires (p.165), skits, small group discussions and whole-class discussions. She got the participants to express their beliefs about language learning and their classroom activities. The approach was not the researcher telling participants how to act in small groups, but the learners sharing their ideas, beliefs and needs. The students found that the skits put on by their teachers illustrated ways they, the students, really behaved, for example: talking in L1, wasting time, answering 'yes' even when they did not understand. The students felt that the skits, and therefore their teachers, showed, not ordered, them how to waste less time and get more out of small group work. She found that the forum helped improve their sense of community and students reported that it modified their behaviour. Ewald's main point stemming from this work was students should be involved in research on pedagogy (cf. Van Lier, 2004) and teachers should develop activities that encourage students to explore issues related to their own learning. Because small group discussions are very common open-ended, opinion-gap tasks, this research provides a model for addressing some of the issues that have been raised by critics of tasks, such as off-task talk, L1 use, telegraphic language (minimal L2 use) (e.g. Carless, 2004; R. Sato, 2010; Seedhouse, 2005; Sheen, 2003). One issue with Ewald's study is that she did not address changes to learner behaviour through observation and classroom recording, only through learner self-report. In addition, she does not report on students whose behaviour did not change after the intervention. This study will address these two elements.

Lam (2009) investigated whether explicitly teaching seven metacognitive strategies to her learners in Hong Kong would improve their small group discussion task performance. She identified: 'problem identification', 'planning content', 'planning language', 'evaluation', 'asking for help', 'giving help', and 'positive self-talk' (p. 131) as the target strategies taught to the treatment group. The control group received no metacognitive strategy instruction (MCSI). Data were collected three times: weeks 1, 10 and 20 and consisted of the learners' performance in group-work discussions, self-report questionnaires, observations of learners' strategy use, and stimulated recall interviews. She found the treatment class typically performed the group discussion task better than the control group based on ratings of task effectiveness and English proficiency. She found that there was an overall increase in strategy use, particularly 'problem identification' (p.143). She argued for the awareness-raising value of the MCSI. She also

found that learners typically deployed a problem identification strategy when the following discussion task involved a prioritization. In response to this finding, she argued that teachers could match tasks with strategy instruction relevant to that task.

Lam's (2009) work differs from Ewald's (2004) study in that strategies were identified and presented to the learners, rather than having the learners' awareness raised implicitly. The disadvantage for teachers of doing things this way is that they need to be instructed in MCSI and need to have (parts of) more than one lesson available for it. However, Lam's (2009) findings show that if students are trained in strategies for task performance, their performance will improve leading to better learning opportunities from the task.

The participants in Lam's (2009) study showed the clearest change in their use of planning strategies, and not interactive ones. To investigate the effect of interactive strategy training, M. Sato and Ballinger (2012) took a mixed-methods approach. They trained learners with strategies to enhance the effect of peer interaction. What is unusual with their work is that they report about two very different studies, from two different approaches. The first, an experimental approach was one in which Japanese university students were taught to give peer feedback, and their later task production output was measured and compared. The other was grounded in sociocultural theory, and was one in which elementary French immersion students in Canada were taught to collaborate. Sato and Ballinger found that the Japanese students improved the frequency of feedback and self-corrections, and the overall accuracy of their production increased after the training. The Canadian students' activity was recorded and analysed for contextual factors that influenced their use of peer feedback. The authors claimed that peer interaction was enhanced by the training, but the outcome of interaction and feedback was dependent on the learners having a "collaborative mindset" (p.173). This echoes findings by Philp et al (2010), who demonstrated that willingness to attend to errors was due to social considerations and not due to cognitive demands of the task. However, Sato and Ballinger (2012) made no attempt to integrate cognitive and SCT findings across the two studies. They leave the two studies separate and fail to investigate if ecological factors in each situation aided in the development of the collaborative mindset. We shall see, in Chapter 8, below, that such factors may well be essential to the formation of such a mindset.

Seifoori and Vahidi (2012) showed the effect of strategy training on the fluency of speech, something not as dependent on the interlocutor as feedback training and collaboration

strategies are. They trained learners to attend to the fluency of their speech during online task planning with the expectation that such training could offset the trade-off effects of increases in fluency decreasing accuracy and vice-versa (see Skehan, 1998, 2009). Their strategies included avoiding disfluencies such as false starts and repetition, and using common fillers (Seifoori & Vahidi, 2012, p. 105-6). They found that the training improved the fluency of the trained speakers without sacrificing accuracy. The study targeted the strategies to be employed online, as the conversation progressed. These were also strategies employed by a speaker during their turn and would be less influenced by other learners.

To sum up this section, research on learner awareness has shown that learners can incorporate explicitly presented strategies into their talk and this incorporation of strategies leads to improved task performance (Lam, 2009; M. Sato & Ballinger, 2012; Seifoori & Vahidi, 2012). Implicit raising of learner awareness also improved learner behaviour in task performance as evidenced by self-reports (Ewald, 2004). The pedagogical advantage of the implicit approach was it took less time and allowed the students to voice their ideas about their learning. It remains to be seen if implicitly raising learner's awareness can be empirically demonstrated to show improved task performance.

2.9 Chapter Summary

This chapter has defined tasks and approaches to using them. It follows Samuda and Bygate's (2008) definition. It reviewed current theoretical positions about how tasks lead to language learning. It discussed research on how task performance has been investigated. The focus throughout the chapter was on research done on tasks in classrooms. I discussed both cognitive and sociocultural approaches because both provide valuable and complementary perspectives on task-based instruction. I then argued for an ecological perspective in the investigation of tasks as this approach treats context as central to investigation. This perspective will accept experimental data collection and analysis methods as long as they are situated elements in a context and thereby offer insight to that context.

I then discussed Willi's (1996) framework of TBLT. I argued that this framework is

flexible and convenient for teachers. I discussed Willis' notions of the relation of native-speaker input and use of CR activities post-task to increase learning of form and I pointed out that this has not been empirically substantiated. Likewise, this framework was developed at the level of the task, or the lesson, but her approach to syllabus design is left up to the teacher. So, although Willis' (1996) framework was originally based on cognitive theoretical constructs, it requires the teacher in context to determine the level and type of task, so an ecological approach could be taken when investigating it.

I then examined three sub-topics relevant to this research: the role of input in TBLT, task repetition, and learner awareness training. I argued that the relationship of input to task performance has not been demonstrated. There is need for research on when, before or after the task, input should be given to provide learners with opportunities to mine words and expressions from the input or to notice a gap in their ability to produce output and learn from that noticing.

Willis (1996a) suggested that repeating a task after the language focus would be helpful. Other research on task repetition was examined and the number and kind of repetitions (procedural or task) that provide the best learning outcomes has not yet been fully established, particularly in classrooms where participants may not be fully engaged.

Learner awareness training may be helpful for task performance, but, particularly with respect to implicit techniques its usefulness remains promising but in need of further empirical support.

This concludes the review of the literature for this study. The next chapter reports the methods, results and discussion of the first phase of this research, an 'exploratory study' carried out to identify tasks that provided strong learning opportunities. Subsequent chapters report on the second, quantitative phase of the research in which elements discussed in this chapter: input, repetition, and awareness, will be manipulated.

CHAPTER THREE: THE EXPLORATORY STUDY

3.1 Introduction

The previous chapter outlined the theoretical frameworks and previous research that informed this study and raised questions about the use of tasks in Asian contexts. This chapter presents the procedures, results and discussion of an exploratory, observational study of ecologically valid task interaction in Korean university classes carried out to determine which tasks and interventions to investigate in the second phase.

The chapter first provides the purpose for this phase, it outlines the methods (participants, materials, and data collection and analysis procedures), then it presents the results and discusses them in terms of theory and practice, finally concluding with the direction for the next phase of the study.

3.2 Purpose of the exploratory study

The first phase of the study was an investigation of how students performed the ecologically valid tasks implemented as part of their first-year English class at a university in Seoul, South Korea. It set out to determine what learning opportunities arose during task performance with the overall aim of identifying tasks that prompted the students to both complete the task and to attend to their language while doing so (Ellis, 2003, p. 8).

To determine successful task completion, the differences between the task-as-workplan and task-in-process (Breen, 1987, 1989) were considered. To determine attention to language, the relative number and focus of language related episodes (LREs) (Swain & Lapkin, 1995, 1998) and patterns of use of L1 (Butler, 2011; Carless, 2004, 2008; Hall & Cook, 2012) were investigated.

The task which best promoted attention to both grammar and vocabulary while still promoting meaningful communication and task completion was then used in the second

phase of the study. The research questions were as follows.

1. Which tasks did students perform in a manner consistent with the task-as-workplan? If task performances were not consistent with the workplan, how did the changes impact learning opportunities?
2. Which tasks promoted students' attention to both grammar and vocabulary during interaction as determined by the number and type of language related episodes (LREs)?
3. In which tasks did students use more L1? How and when was it used?

3.3 Methods

3.3.1 The research context

The study was carried out in the Practical English Department at a 4-year university located in Seoul, South Korea, in the first semester of the academic year (March – June 2012). Practical English was an English for Academic Purposes (EAP) class that used a Sustained Content Language Teaching (SCLT) approach (Brinton, 2003; Murphy & Stoller, 2001) to prepare students with the language and skills they would need to participate effectively in English-medium courses offered at the university. In general, SCLT simulates regular university courses by presenting one main topic throughout the course, with language instruction arising from this topic. Language learning is enabled by drawing learners' attention to features of the input, through learner-learner interaction, and through explicit instruction and awareness raising activities (Brinton, 2003, pp. 205–209). The focus of this programme was on: academic listening, the taking of notes during lectures and on the discussion of lecture content. The content was American Studies. The textbook *Academic Listening Encounters: American Studies* (Sanabria & Sanabria, 2008) was its core material.

For this particular program, there was a great deal of flexibility afforded instructors by the coordinator, in that although a common approach (SCLT) and common textbooks were assigned, the particular chapters and activities from it to be used, supplementary materials, methodology, and assessment were at the discretion of the instructors.

The first semester of this study began at the end of February 2012 and ended at the end of June 2012, consisting of 16 weeks, with 3 hours of class each week, divided into a 1-hour and a 2-hour lesson on different days.

Prior to beginning the project, I briefed the program director about the aims and goals of the study and received written permission to carry out the study. The president of this university actively encourages research into innovative teaching methods and therefore the programme director was eager to assist.

3.3.2 Participants

3.3.2.1 Student-participants

The participants in the study were students enrolled at the university who had signed up for one of two classes with me. Class A met Mondays at 3pm, Wednesdays from 5-7pm; Class B met Mondays at 4pm, Wednesdays at 3-5pm. They studied a variety of majors, but most studied other foreign languages (e.g. Thai, Vietnamese, Hindi, etc.) or business or politics-related subjects (e.g. business administration, international economics, etc.) Students majoring in English were not permitted into Practical English classes.

Students were informed of the research at the start of the first lesson, to give them a chance to rearrange their schedules if they desired, and again at the start of the second lesson. They were given an information sheet and consent form (translated into Korean) and asked to indicate yes or no (See Appendixes 5 and 6 for the sheets). Only one student, chose not to participate, her data was replaced in the data set with 'XX'.

This may seem like a low class size for an Asian university (cf. McDonough & Chaikitmongkol, 2007 for example), but the university capped enrolment at 30 per class, and there were four other sections of the same course opened at the same time leaving all four sections slightly less than full.

Table 3.1, below summarizes the composition of the participants by gender, year in the university, major and English learning experience. It shows that the two classes were similar in terms of these very general characteristics. The age range for both classes was 20-22 years old.

Table 3.1. Composition of classes A and B

	Class A	Class B
Male	8	7
Female	11	13
First year	18	20
Upper years	1	0
Non-participating students	0	1
TOTAL	19	19

3.3.2.2 Human ethics

I served as both researcher and classroom instructor for the study. This means that the participants were both my students and my collaborators in all aspects of this research. This allowed me to engage with them as both a helper and instructor and as someone who needed their assistance. The relationship of teacher-researcher (me) and the student-participant led to two main concerns.

The first concern was ethical². I received written approval to do the study from the university as well as from the Human Ethics Committee at Victoria University of Wellington. I gave a Korean-language translation of the information sheet to the participants and gave them the option of saying ‘no’ I could not use their words as part of this study (See Appendixes 4 and 5). Only one student took this option. However, students might feel pressured to comply. So I pointed out to them that I taught other classes and had many students, so if they did not wish to participate, there would still be more than enough data. The students also had the option of taking another section of the course as there were four simultaneous sections taught by four different instructors, none of which was full.

Another issue of teacher-led research concerns the potential for observer bias by the researcher and for response bias by the participants. These are genuine concerns with all observational research (Creswell, 2002, 2007; Nunan, 1992; Richards, 2003), and the countermeasures are the same: triangulate data from multiple sources; try to use low-

² This study received Human Ethics Research approval from Victoria University, RM#18762; the information sheets and informed consent forms, both English and Korean can be found in the appendix.

inference measures when possible; use recordings when possible; reflect on decisions made; follow member-checking protocols when needed; and have another independent rater check at least a portion of your data analysis. These were all used and reported in the data analysis.

3.3.3 Data collection procedures

Data was collected during class by video-recording the class, audio recording each small team of participants, and by taking observation notes. Each lesson, prior to class, I set up a single video camera in a position from which most students could be seen. At the start of the lesson, I randomly assigned students into new teams of three to five students each depending on the lesson plan and student numbers. Once they were in their new teams, I placed an audio recorder at each. In my observation notes, I filled in a seating chart and kept track of student movement from team-to-team to aid identification of the speakers when transcribing. In addition to in-class data, at the end of the semester, student-volunteers were recruited for end-of-study interviews. So, the data set used in this phase included over 36 hours of audio recordings of task performances, observation notes, and video recordings along with interview transcripts. Each data source is described below.

Audio Recordings: near the start of each lesson, after students had been assigned into teams, a small voice recorder (model, Sony ICD-PX312F) was placed in the centre of each team's desk. These recorders were numbered and recorder 1 was placed with team 1, recorder 2 with team 2, etc. so the voices on the recordings could then be matched with group members using the seating chart. This model of recorder has microphone settings for certain target ranges. The shortest range, 'voice notes', allowed me to hear the students' voices clearly while eliminating background noise from other groups.

Video recording: a single video camera was used to record general classroom activity. Typically it was placed at the back of the room so I could use the video to confirm my observation notes about what activity was being done at what time. In some lessons, the camera was situated near a single group to capture their interaction more closely.

Teacher's notes: I kept a notebook with me at all times to note down seating arrangements, approximate start and finish times of each activity, student behaviour and roles, common errors I overheard and student language that needed following up, along with other thoughts about the classroom as they arose.

Lesson plans: Prior to each lesson, a lesson plan was made and placed in a folder that was referred to during class. It made a permanent record of the task-as-workplan (Breen, 1987, 1989) from which comparisons could be made to the task-as-process under study.

Interviews: At the end of the semester, students were invited to participate in a one-hour stimulated recall interview with the researcher. During this interview, excerpts of their in-task talk were played for them and students were invited to comment on their performance, on the class in general and were invited to ask me any questions they had. The interview was semi-structured, and the protocol is in Appendix 5, below.

3.3.4 Data analysis procedures

After each class, data was saved to computer for analysis. I then read over the lesson plan, observation notes and watched the complete video in order to make an outline of classroom events and timings. Using the video and notes I chose one group for further analysis. The basis of the choice varied; perhaps the group's performance was different from the expected performance in some way, I overheard something in their in-task talk that interested me as an instructor or the group may have had some members whose talk I had not yet examined in detail.

Next, I transcribed the task performance. This included the talk of any member of the group who moved to another group to report their group's ideas. Voicewalker (Du Bois, Corston-Oliver, Holton, & Norris, 1998), free software designed for linguistics research that automatically steps audio files forward five seconds, back four, then forward again, was used to aid the transcribing. Transcribing of L1 Korean was challenging at times, as my Korean proficiency is high beginner. To compensate, I asked a Korean professor at the university to double-check my transcription, and when necessary, I played the recording for the participating students and asked them to confirm what had been translated. This last option was only done after the end of the semester, so as not to influence the study.

During and after the transcribing process, patterns of interaction were identified and classified. I was looking for evidence of learning, as shown by LREs, for how the tasks were being approached by the participants (i.e. the relation of the task performance to the workplan) and for L1 use. After transcriptions and preliminary coding was done, the recording was listened to again to confirm the contents of the transcripts. Then other groups' recordings were listened to in order to determine if the group being focused on

was typical or atypical.

I then wrote a short report about the task. At the end of the semester, four volunteers came to my office for the stimulated-recall interview. They read a preliminary transcript of their task performance and listened to their talk. It is normal procedure to do such an interview as soon as possible after the event (Mackey & Gass, 2005, p. 78), but doing so might have distorted students' in-class behaviour. The interview started by me asking questions from the interview protocol but as the interview was semi-structured, follow-up questions and deviations from the set questions were permitted. The students were invited to challenge what I had written about their transcripts. The interviews were also recorded and analysed for their comments as part of the data triangulation.

3.3.5 Materials: The assigned textbook and supplementary tasks

3.3.5.1 The assigned textbook

The primary material for the course was the course textbook (Sanabria & Sanabria, 2008). This book's aim is to provide ESL learners the academic listening and note taking skills needed to participate in tertiary education in the USA. The programme director felt this textbook was appropriate for the programme (Sa, personal communication) because the students would have to take English-medium courses in their majors. It was designed for a 70-hour course, and there were only 48 hours of instruction, so large chunks of the textbook were skipped by all instructors.

3.3.5.2 Supplementary tasks

Considerable leeway was given for instructors to supplement the text with their own materials. I chose tasks that would allow learners the opportunity to integrate what was in the text with their own ideas and opinions. I used the same tasks I had used when teaching the programme in previous years, tasks that had engaged student interest. Willis and Willis say that "Tasks are not a substitute for interesting topics which engage learners' interest, but they can enhance that engagement and interest" (Willis & Willis, 2007, p. 11). The textbook provided the topic, and the tasks were chosen for how they were expected to enhance learners' engagement with it. Six general task types were used to supplement the textbook's activities: listing, discussion (sharing personal opinion, information or experiences), creative role plays, values clarification, and creative simulations. The characteristics of these task types are outlined in Table 3.2. (A full description of these

tasks' workplans is provided in Appendix 8.)

Table 3.2. Characteristics of the tasks used

Task type	Interaction type	Response type	Information flow	Expected Outcome	Expected behaviour in each phase of the task cycle (J. Willis, 1996b)		
					Activity	Planning	Report
Listing	convergent	open	all-way	a written list and rationale for including items in the list	List the items. Discuss and explain reasons for suggesting an item's inclusion in the list	Write a final version of the list, noting the rationale for including each item	Read the list to the class along with rationale for including items in the list
Discussion (Sharing personal opinion/information)	divergent	open	all-way	a summary of the discussion	Ask and answer questions on the topic	Summarise the main ideas of their talk	Present the summary to the class
Creative role play	convergent	open	all-way	a performance of the role play at the front of the class	Make a script related to a given scenario	Practice acting the script	Act out the scenario at the front of the class
Jigsaw listening	convergent	closed	one-way	a retold lecture	One partner listens, takes notes and then retells the lecture contents	Both partners work together to complete a graphic organiser	The second partner retells the lecture to the class (using the graphic organiser)
Creative simulation	convergent	open	all-way	a presentation of the item the team created to another team (in competition with a third team)	Create an item to sell or get funding for; then create a sales presentation	Practice the sales presentation	Present your team's item and compete with another team to get funding
Values clarification	convergent	restricted open	all-way	a written revision of a controversial statement	Discuss the controversial statement	Re-write the statement in a mutually acceptable way	Present your revised statements and reasons for the revision to the class

Based on Ellis (2003) and Willis (1996) and my lesson plans

No consideration was given *a priori* to the number of times a task type was used, or its relation to other tasks used in different lessons. The sole consideration for use was the topic of the course material and how to best engage learners with that topic. As a result, the implementation of task type emerged ecologically as the course progressed based on my observations of task performance. The resulting tasks used across the 32 lessons of the course are shown in Table 3.3. This table omits lessons in which students did not do a supplementary task.

Table 3.3. Schedule of tasks, topic and textbook chapters

Lesson number	Task type	Topic	Textbook chapter.section
2	Listing	Questions for the teacher	N/A
3	Sharing personal opinion/information	Introducing self/others	N/A
4	Listing	National problems	1.2
	Creative role play	Voting and elections	1.2
5	Values clarification	Censorship	2.2
7	Sharing personal information/opinion	Civil Rights	2.4
9	Sharing personal information/opinion	Language learning	3.1
10	Listing	Government agencies	3.4
	Creative simulation	NGO funding	3.4
11	Sharing personal information/opinion	Immigration	3.1
12	Creative role play	Immigration	3.2
17	Sharing personal experience/opinion	Immigration	3.4
18	Jigsaw listening	Civil Rights	5.4
19	Sharing personal experience/opinion	Discrimination	5.4
20	Creative simulation	Superheroes	7.2
21	Values clarification	Workplace	8.2
23	Jigsaw listening	Folk heroes	8.4
24	Listing	New inventions	9.2
	Creative simulation	New invention	9.2
25	Jigsaw listening	American music	9.4
26	Listing	Healthy living	10.2
	Creative simulation	Campus health	10.2

Of the 32 lessons, 18 involved the use of a supplementary task. The other lessons did not. In these lessons, either the textbook was used without supplementation or there were tests, national holidays or university events that led to cancellation of the class.

3.4 Results and discussion of the exploratory phase

The results section examines how students performed each of the task types. It starts with general observations and then provides evidence from the learners' talk. Finally, the chapter connects the findings to other research and provides a rationale for choosing one task to study in further detail in the next phase of the research.

3.4.1 Task performance

This section provides an overall picture of how the tasks were carried out by the students. It compares the task as workplan, shown in Table 3.2 (above) and the observed task performance, shown in Table 3.4 (below).

First, the data revealed that students always performed the tasks, that is, they reached the expected outcome listed in Tables 3.2 and 3.4. There was no case of a team not finishing a task. However, the tasks were not always performed in accordance with the workplan. Table 3.4 provides an overview of the observed patterns of behaviour in each phase of the task cycle (J. Willis, 1996b). Comparing Table 3.4 with Table 3.2 revealed the adjustments to the workplan made by students as they performed their tasks (cf. Breen, 1987). The principal adjustment made by the students was to use the planning phase as the opportunity to add more content. Details of how this influenced learning opportunities are described following Table 3.4.

Table 3.4. Summary table of observed task performance

Task type	Expected Outcome	Observed patterns of behaviour in each phase of the task cycle (J. Willis, 1996b)		
		Activity	Planning	Report
Listing	a written list and rationale for including items in the list	As students listed the items as required they discussed and evaluated them. Talk was generally on-topic. LREs typically referred to vocabulary and often involved L1.	Students used planning time to add more items. In cases when students did plan their report, grammatical LREs were found.	Students reported the items listed, however, they needed prompting to provide reasons for including their listed items otherwise they frequently failed to provide them.
Discussion (Sharing personal opinion/information)	a summary of the discussion	Students often went off topic and often used L1. There were very few LREs and they were only vocabulary. Teams rarely took notes during the discussion	The lack of notes led to difficulty summarizing their discussion	Students were unclear what to include in a summary of the report. Often, the presenter's opinion was what was reported, not the group's discussion.
Creative role play	a performance of the role play at the front of the class	Students worked together to make a script. L1 was used for idea generation as well as a language resource. LREs were typically related to vocabulary.	Few teams practiced performing their script. After being told to practice, they typically added more content. L1 use sometimes increased at this time.	Scripts were generally a series of sentences read aloud in turn at the front of the room. Often it was written on a single sheet of paper and shared.
Jigsaw listening	a retold lecture	Notes were not well taken, so the retelling by the first partner to the second partner was often factually incorrect. There was a great deal of L1 used to verify facts from the listening. The graphic organisers were typically completed properly only when pairs were combined into teams of four students	Planning time was generally used to add more content or clarify what was noted on the graphic organisers.	Students had great difficulty presenting a retold version of the lecture to the class. They used the graphic organiser to aid in presenting, but items on it were not elaborated with detail from the recording.

Creative simulation	a presentation of the item the team created to another team (in competition with a third team)	Students worked together to create an item to sell to other groups. There was a great deal of L1 used, and the amount of L1 was observed to increase each time the task was done (i.e. the 3 rd simulation had the most L1 use). L1 was used for ideas and for language resources, as well as for social talk.	Teams did not rehearse the presentation at this time. Typically planning time was used to add more content to the report. Teams used much more L1 at this time to add new content. Even teams that had been using mostly L2, typically switched to L1. A team would only rehearse the presentation when the teacher instructed them to rehearse and he stood nearby to watch them practice.	Because they had not rehearsed the presentation, team members who were presenting to new teams often had to spontaneously come up with English versions of their Korean-sourced presentations. Presenters resorted to telegraphic language, gesture and L1 to explain their ideas.
Values clarification	a written revision of a controversial statement	Students revised the prompts in one of two ways. Either each team member stated their level of agreement and then discussed their reasons in turn. Then they worked on revisions. Or one team member started to revise the prompt and as revisions were suggested, they explained their choice of words based on their opinions. There were often vocabulary-related LREs at the start of the talk about a prompt. Students used L1 mostly in LREs.	Students wrote down the revised version of their prompts. As they wrote there were often grammar-related LREs. Presenters practiced their report and team members suggested changes. Content-relevant L1 and LREs happened in this stage, too.	The presentations included the vocabulary of the prompts and the reasons they had discussed either in the Activity or Planning stages.

3.4.1.1 Failure to plan or rehearse the report

Willis (1996) argue that learners push their output to a more formally accurate version during the planning stage, so the teams were given time to practice their final reports before presenting them to the class (or to another team). In order to make a clear distinction between activity and planning stages, I stopped their activity, reminded them of what form the report would take, and announced to the class that they needed to prepare their presentation and rehearse it at their teams before presenting it to the class. In short, the planning stage instructions required students to organise their ideas and then rehearse how they would be presented to the class.

It was observed, however, that students typically used the planning phase to add more content, and with the added time pressure created by the upcoming deadline, students often switched to L1 to do this. They rarely practiced presenting the report verbatim prior to presenting it or negotiated the accuracy of the language in the report. So, despite clear instructions to use the planning phase to improve the quality of the language of their final report, something Willis (1996) claimed would happen spontaneously, the planning phase was used to improve the quality of the report's propositional content and almost no attention was devoted to linguistic accuracy. The lack of planning led to spontaneous presentations characterised by: disfluencies, telegraphic speech or, in the case of role play, reading.

Extract 3.1 shows the planning talk of a creative role-play from Lesson 24. The team members are preparing to present their new (imaginary) invention, a personal body scanner to help you order clothes that are the right size when shopping online. They were preparing for a simulation of a sales presentation, in which one member of their team would present their item to a new team for one minute, and then compete with members of other teams to sell their item. Before this extract begins, the instructor had told the class, “*You can't bring your papers [i.e. notes about the invention] with you, so practice, practice, practice.*”

Extract 3.1. Planning in creative role play

Turn	Speaker	Utterance (what they said)
1	Amy	L1= <i>neomujoh-a</i> 너무 좋아 (tr=I really like it.) I like it.
2	Teacher	so who is your seller? Who's captain?

3	Amy	//captain.\\ (gestures to herself)
4	Doug	//she's captain \\
5	Teacher	so who's your sellers? It's Your choice. (T walks away)
6	Doug	L1 <i>geulaeseo eolmalago</i> 그래서얼마라고? (tr=So how much is it?)
7	Amy	L1 <i>chogibiyong 500dalleo, eobdeiteuhalttaemada 10dalleo</i> 초기비용 500달러, 업데이트할때마다 10달러. (tr=Initial price is \$500 and \$10 for updates.)
8	Stephanie	L1 <i>tledon</i> 떼돈 (tr=big money)
9	Doug	L1 <i>gagyeogmyeon-eseoma-eum-eandeul-eohalgeogat-a.</i> <i>500dalleonanaegosalamdeul-i</i> 가격면에서마음에안들어할거같아. 500달러나내고사람들이.....? (tr=People won't like the price. Will they pay 500 dollars...?)
10	Amy	L1 <i>salamdeul-eunkamelamansamyeondwae</i> 사람들은카메라만사면돼. (tr=They only need to buy cameras.)
11	Cindy	L1 <i>sajin-eul-ollimyeongeunyangjeojeollo 3D gadoeneun</i> 사진을올리면그냥저절로 3D가되는...(tr=photos automatically changes to 3D when they upload them.)
12	Amy	L1 <i>ya potosyab do geunyang-gongjijaloda daunbadjanh-a. geuleongeoscheoleom</i> <i>gieob-eseoneun don-eul jugosayadwae. Geunde gaeindeul-eun</i> <i>bulbeobdaunlodeuleul haneungeoji</i> 야포토샵도그냥공짜로다다운받잖아. 그런것처럼기업에서는돈을주고사야 돼. 근데개인들은불법다운로드를하는거지. (tr=Hey, people just download Photoshop for free. Like that, companies have to pay for it, but people illegally download it.) L1 <i>man-yag-e nuga neomu bissadago hamyeon geu yaegileul haejwo</i> 만약에누가너무비싸다고하면그얘기를해줘. (tr=if someone says it's too expensive, tell it.)
13	Cindy	L1 <i>eo? jinjjayo?</i> 어? 진짜요? (tr= Oh, really?)
14	Doug	L1 <i>nan chalali daleunsaeng-gag-eulhaessneunde. Aigeo ileohge jinjihageyae</i> <i>gihaedodoena, igeoleul? Geunyang magjijeom gat-eungeoleul hana jeonghaeseo</i> <i>salam ileohge ttagijiggo aidigat-eungeoleul mandeul-eoseo</i> 난차라리다른생각을했는데. 아이거이렇게진지하게얘기해도되나, 이거를? 그냥막지점같은거를하나정해서사람이렇게딱찍고아이디같은거를만들어 서..... (tr=I have completely different thought. Ah, is it okay to say it seriously like this? Just make a branch or something ,take a picture and make an ID....)
15	Amy	L1 <i>seutikeosajin cheoleom</i> 스티커사진처럼? (like sticker photo?)
16	Doug	You like...
17	Teacher	L1 <i>yeong-eolo</i> 영어로! (in English!)
18	Doug	there is my own body size . you understand ?

19	Stephanie	yeah, I understand.
20	Doug	L1 <i>chalali</i> 차라리 (Rather)
21	Amy	L1 <i>chalali</i> 차라리(Rather)
22	Doug	L1 <i>chalali ... chalaliga yeong-eolo mwoya?</i> 차라리...차라리가 영어로 뭐야? (Rather.... What is chalali in English?)
23	Amy	L1 <i>museun jogeu manbagseuedeul-eogaiss-eumyeonmom-eu ljjwaghaegajigo geuge meinbodeulo jeonsong-idwaeseo</i> 무슨조그만박스에 들어가있으면몸을쭉해가지고그게메인보드로전송이돼서... (tr=If you enter something like a tiny box, it scans your body and sends to main board...)
24	Doug	Yes, yes.... L1 <i>mwoyeossji? sijang-ui suyowagong-geub-e uihae ..gyun hyeong jeom eseottag-ileohge</i> 뭐였지? 시장의수요와공급에의해..균형점에서딱이렇게... (tr= What was that? By demand and supply of the market... at the equilibrium....like...)
25	Amy	in English...L1 <i>suyo</i> 수요? (tr=demand)
26	Doug	demand and supply! don't you know ? don't you know demand and supply? what's this...this point? (gesture?)
27	Cindy	Equilibrium.
28	Doug	L1 <i>amajda</i> 아맞다. (tr=Right.) Equilibrium.
29	Amy	oh. smart.
30	Doug	equilibrium.
31	Cindy	haha.
32	Doug	demand, supply, equilibrium . it's the price.
33	Amy	okay, you're a good seller. break time!
34	Stephanie	L1 <i>na jinja seller ya</i> 나진짜 seller야? (tr= Am I a seller?)
35	Amy	L1 <i>du myeong-ieyo</i> 두명이에요 (tr= Two people)
36	Doug	L1 <i>ileohge nanwo yaji</i> 이렇게나눠야지. (tr=we should divide like this.) L1 <i>yeogidatto sad story-leul seokk-eo yagessda</i> 여기다또 sad story를섞어야겠다. (tr= I will put a sad story in it.)
37	Cindy	sad story ? haha.
38	Amy	L1 <i>nuga sal-I neomu manh-ijjyoseoos-eul mos-saib-eungeoya</i> 누가살이너무많이꺼서옷을못사입은거야.(tr= Someone wasn't able to buy some clothes because she was too fat) sad and true story.
39	Teacher	okay, that's plenty of time . MOVers, please stand up.

The extract exemplifies most teams' planning. The entire talk was concerned with

finishing the creation of the invention. At turn 6, Doug started by asking the price and upon hearing it, he commented that that no one would pay that much. The price eventually was revised by turn 12. Additional features of the product were discussed throughout the extract. However, the participants who eventually made the presentations, the ‘sellers’, never practiced what they were to say to the other groups as instructed by the teacher. Throughout the time when the task-as-workplan called for students to rehearse the sale, the students added more detail to their invention. In addition, they did this almost exclusively with their L1 (Korean) – a point I return to in section 3.4.3. This meant that when Stephanie and Doug separately moved to new teams (Stephanie to Team 2, and Doug to Team 5) to sell their invention, they were doing the presentation for the first time. Extracts 3.2 and 3.3 are transcripts of these presentations.

Extract 3.2 Stephanie’s presentation at Team 2

Mmmm Our... (cough) new technology's name is Fitting Model. Ah...ah....First you take a picture of your body in 3D camera and put it on- online and ah... you go an- a online shopping home page and you click the clothes you want and it will show you how it looks...like...m... if it's big it will change the size or repair the length and if it's small it will tell you which part you need to go on a diet. The cost is...five hundred dollars and the online shopping.... web sites owners need to buy it but...and... users just need to pay ten dollars when it updates

Extract 3.3. Doug’s presentation at Team 5

Our device is fitting model...um..it's like..if this devices take a photo of our body, and check the size and fit with the clothes which is on the Internet. So you can buy clothes without going stores. Fantastic?

The task-as-workplan held that the initial presentations would be nearly equivalent, but the above extracts reveal that Stephanie’s was much longer and more detailed than Doug’s was. They were supposed to have spoken for a minute, and the instructor had required each team to time the presentations as they listened. Doug did not reach half of the minimum time. It is probable that if he had heard Stephanie’s presentation, and if his team had heard his presentation in this form, Stephanie’s model (cf. Kim & McDonough, 2011) and peer feedback from his team members would have helped him make it longer and more detailed.

Later, as Doug tried to convince the listeners at Team 5 to choose his invention, he took a strategic approach to finishing the task. For example, when one of his listeners said his item was “*convenient*”, Doug immediately attempted to finish the presentation by saying, “*Yeah! Fantastic! I sold it. You agree with me, right?*” His listeners did not immediately agree and there was more discussion, however Doug never mentioned the price, the 3D image, or who pays for using the device. Planning and rehearsing in advance might have improved Doug’s responses to challenges from his listeners (cf. Bygate, 2001).

The creative simulation tasks were not the only task types in which the students chose not to plan but to continue the task’s activity. In creative role plays, teams only rehearsed their scripts when the instructor specifically told them to and stood nearby to watch. Otherwise they invariably used the time to add more lines of dialogue. This meant that resulting performance of the role play at the front of the classroom involved the team handing a single sheet of paper back and forth and reading their dialogue aloud, with no attempt made to learn or remember the lines.

This reading aloud of the final report limited the learning opportunity from the task. The intention of the task was to create and remember a short (less than one minute per speaker) dialogue originating in a scenario relevant to their textbook’s topic. Learning opportunities from interactive writing remained unchanged. However, reading aloud negated the language learning benefits of memorising even a short dialogue (cf. Wray & Fitzpatrick, 2008).

In discussion tasks, teams were to summarise their discussion and one member was to present it to the class. However, teams rarely did this. Instead their presenter gave his or her opinions that they had said previously to their team, regardless of whether the other members agreed with them or not. This tactic meant that during the discussion, team members did not have to concentrate on what the others said or integrate their ideas and language into the final report, thereby limiting potential learning opportunities.

Space limitations preclude a detailed examination of all the task types, however, in general, all the tasks suffered from students changing the planning time to further idea-generation time. The values clarification task suffered least because it required the re-writing of a prompt to a mutually agreeable form, thereby necessitating discussion and integration of all team members’ ideas. The presenter could remember what had been said and include it in the final report with limited further planning.

3.4.1.2 Failure to understand the input

The one task that was heavily dependent on pre-task input was the jigsaw task. It involved pairs of students each listening to a recording of half of a lecture from their textbook and taking notes on a blank sheet of paper. They then used the notes to retell the lecture to their partner, who completed a graphic organiser in their textbook (e.g. Sanabria & Sanabria, 2008, pp. 78–9). Once both partners had listened and shared notes, they combined with another pair to compare and confirm their answers. They then used the graphic organiser to present that part of the lecture to the class. The main issue for task performance was not the students approach but the difficulty of the input.

Extract 3.4 shows one team struggling to complete the graphic organiser on the topic of the Civil Rights movement in the USA (from Lesson 18). The extract refers to the story of Rosa Parks who refused to give up her seat on a bus. The pair of William and Laura joined Tiffany and Sunny for their final plan before Laura will present to the class.

Extract 3.4. Jigsaw task retelling

Turn	Speaker	Utterance
11	William	lisa parks- rosa rosa parks
12	Sunny	rosa parks
13	William	rosa parade sit in the white person area
14	Tiffany	lives in whi- white person-
15	William	no no no seat!
16	Tiffany	seat ?
17	Laura	ah: she or he
18	William	she
19	Laura	she cannot sit the seats that
20	William	white girl
21	Laura	she sits on white area?
22	William	uh that uh no no
23	Laura	and then she struggles ?
24	Tiffany	sorry
25	William	no that white person's area that (? :mm) uh only only white person (? :mm) can
26	Tiffany	sit
27	William	sit in the bus
28	Laura	ah
29	William	uh huh
30	Laura	cannot sit in the bus
31	William	uh lisa- rosa uh she is black woman (? : mm) lisa rosa sit in the white person area that only white person can sit in the bus

It is clear from the transcript that the team has not understood that Rosa Parks was sitting in an area for blacks on the bus and refused to give up her seat to a white person when the

white seating area was full. Laura's presentation to the class is in Extract 3.5.

Extract 3.5. Laura's report to the class

beginning of that of that uh civil right is is because uh Rosem- Rose Parks h um uh he- she went to she went to uh bus and uh and there is separate for two areas h* it's for black people and it's for whi- white people and uh uh there is that it uh that is uh uh laws in in that in that time in that time of in that time of America and uh and even though uh restaurants has this kind of difference um after that um more and more conte- contest and demonstrations to uh to to fight for the equal rights h* and the achievement is is that other groups stills to s-sti- still st-st- still stand up up to fight for rights*

Extract 3.5 shows Laura's final presentation to the class. Unsurprisingly, she has missed most of the details of Rosa Parks, and most other elements of the lecture. After the jigsaw task ended, students asked for the script of the lecture which was not included in their textbook (Sanabria & Sanabria, 2008), but was in the teacher's guide. Because they asked, I showed it on the screen at the front of the room and used a highlighter to follow the words as the recording played. I noted in my observations at the time that students reacted very positively to seeing the script and hearing the talk simultaneously.

I also recorded in my notes, that the students' difficulty with the language of the task, language the administration had required them to master, might be overcome by having them do post-task consciousness raising activities on the lecture script as urged by Willis (J. Willis & Willis, 2007; J. Willis, 1996b). This became a theme of the second phase of this study (Chapter 4 and 6).

To sum up, the key observed differences in task performance vs. workplan were: (1) students did not plan language, but rather used the planning phase to add content, and (2) in the jigsaw task, task performance was hindered by the difficulty of the input in their textbook, students needed, and requested, the written version.

3.4.2 Attention to grammar and meaning

The planning phase of the task cycle was proposed by Willis (1996) to be the locus of change from spontaneous, informal speech to more permanent, formal, public speech. It was predicted that this part of the task would feature language related episodes (LREs) related to both grammar and vocabulary. However, students did not use this time to plan,

but rather they used it to add content, often using L1 (see the next subsection). This leads to the question then, of, if any of the tasks promoted focus on meaning and form as evidenced by grammatical and vocabulary language related episodes (LREs) during the task activity phase (per J. Willis, 1996b).

We saw in Extract 3.1 two vocabulary-related language related episodes (LREv's) for *rather*, *demand* and *equilibrium* at turns 22, 26 and 27 respectively. LREv's were relatively common in all tasks. As many of these were mediated through L1 (as Extract 3.1's examples were), it suggests that teachers should not prohibit L1 use in the classroom (more on this in the next subsection) as students use L1 to explicitly resource their L2 as suggested by Swain's notion of languaging (Swain, 2006, 2010). What was not seen in Extract 3.1 was the presence of grammatical LREs.

There were two tasks where LREg were consistently observed: the values clarification task and listing tasks – tasks which required sentences/phrases to be listed. There was one task where LREg were predicted and had few, the listening jigsaw task.

3.4.2.1 Values clarification task

The value clarification tasks generated the most engagement with grammar as well as vocabulary, as manifested by LREs. Extract 3.6 is from the values clarification task done in Lesson 21 of the semester. The students were discussing and then revising statements on the subject of child rearing and discipline. This team, consisting of Sunny, Harold and Crystal, had discussed the item “*It should be against the law to spank children*,” and in the extract they were considering how to rewrite it to a form that all three of them would strongly agree with.

Extract 3.6. LREg co-occurring with writing during a values clarification task

Turn	Speaker	Utterance
135	Harold	uh how how do we change it should-
136	Crystal	-it should be against the law to spank children (reading) it should be against the law to . extreme physical punishment ? to children . (Harold: mm) extreme one should be punished
137	Harold	okay so it should be against to-law the law to -
138	Crystal	//extreme\\
139	Harold	//abuse children\\
140	Sunny	abuse ? abuse
141	Harold	L1 <i>hagdaehada</i> 학대하다 (tr =maltreat)
142	Sunny	ah abuse
143	Crystal	like extreme physical punishment ?

144	Harold	extreme physical punishment –how do we make that sentence
145	Crystal	it should be against the law to extreme physical punishment to children
146	Sunny	(unint.)
147	Harold	ah
148	Crystal	extreme physical
149	Harold	not extremely physical punishment ? ex-extremely physically
150	Crystal	but what what is the extreme extremely physical punishment like um h
151	Sunny	it is vague
152	Crystal	it is vague um
153	Harold	well to
154	Sunny	because you know (2.0)
155	Crystal	uh uh like hit the child for like four hours (Sunny: ah !) without stopping hh
156	Sunny	(3.0) okay okay okay
157	Harold	extremely physically punishment //extremely \\\
158	T	(to the class, but this team keeps talking) //okay we have to stop now (whistles) \\\
159	Crystal	(to Harold while T is speaking) //extremely extreme physical punishment\\ extreme physical punishment
160	T	lovely captains //your job is NOW \\\ is to point at one member of your team and say YOU
161	Harold	but //punishment is L1 <i>myeong</i> 명(tr=noun) \\\ law to it is been law to and to it-
162	Crystal	-and extreme is adjective
163	T	(continuing) AND ? also TELL them which sentence one two three four five or six
164	Harold	//and to to and verb verb should be came . after to\\
165	Crystal	uh I don't know about grammar

Extract 3.6 shows that this team's talk included a push toward more accurate form as they revised the prompt. There were two grammatical LREs in this one short stretch of talk. First, Crystal had made a revised version of the statement at turn 145, "*it should be against the law to extreme physical punishment to children.*" At turn 149, Harold noticed an issue with her grammar, and he asked "*not extremely physical punishment?*" Crystal self-corrected *extreme* to *extremely* in the next turn, and continued trying to clarify what *extremely physical punishment* would be indicating that she had noticed Harold's point about the grammar and changed her own. Later, however, she changed it back to, "*extreme physical punishment*" (turn 159). She and Harold continued to discuss the correct form while the teacher was giving instructions. They even used metalinguistic terms for grammar items, *noun*, *adjective* and *verb* to express themselves (turns 161, 162 and 164). The second LREg occurred at Turn 164. Harold again questioned what Crystal had said, focusing on the to-infinitive construction, which she had got incorrect in turn 145.

The talk in Extract 3.6 was motivated by the demand to write a new version of the prompt.

My observation notes indicated that Harold was his team's writer and the transcript revealed that he was the member focusing on the grammatical accuracy of their re-written prompt. This suggests a possible relationship between collaborative writing and the focus on grammatical accuracy.

3.4.2.2 Listing tasks

This relation between interactive writing and a push for accuracy, as represented by LREg's in the transcript, was observed in some listing tasks as well. When the task required students to list ideas and not words, LREg appeared as the participants wrote the ideas down. Extract 3.7 is part of a listing task in which students were asked to list things they want the next national government to do. It shows four students trying to generate the item, "*The government should make a website where cultural contents can be downloaded at a low price.*"

Extract 3.7. LREg co-occurring with writing in a listing task

Turn	Speaker	Utterance
124	Olivia	download homepage?
125	Julia	uhm government make homepage for low fee (Harold: hmm, Olivia: hmm)
126	Sunny	yes
127	Olivia	you like it
128	Sunny	really
129	Julia	it's really new thing
130	Sunny	such
131	Harold	government should support the music download fee (laughter)
132	Julia	music-
133	Olivia	-homepage
134	Julia	movie culture
135	Harold	//L1 eotteohge 어떻게 ? (tr= how?) \\ how can I write? government made music download
136	Sunny	//make\\
137	Olivia	site site
138	Sunny	make
139	Olivia	support
140	Harold	which gives lower-lower price
141	Olivia	good
142	Sunny	good – captain
143	Julia	and not the song . ah (others laugh)
144	Harold	what?
145	Julia	mo-more material
146	Olivia	yes yes
147	Julia	culture material (Harold laughs)
148	Olivia	culture?
149	Sunny	overall culture area such as music movie
150	Olivia	overall culture homepage (laughs)

151	Harold	// L1 <i>otteohge</i> 어떻게 ? (tr= how?) \ (laughs) how can I write?
152	Sunny	culturally
153	Harold	huh?
154	Sunny	cultural
155	Harold	government should-uh support cu-cultural cultural cultural
156	Olivia	make-uh site-uh
157	Harold	government made
158	Sunny	make site
159	Harold	ma:ke
160	Olivia	government makes
161	Harold	site-
162	Olivia	-site- for
163	Harold	which
164	Olivia	government will make ah .
165	Harold	//government make site\\ which gives ah
166	Olivia	//government make site\\
167	Julia	made
168	Sunny	make
169	Julia	make
170	Sunny	make
171	Olivia	made no
172	Harold	ah government should make

The LREg's were about the correct form of '*make*' and '*culture*'. Twice Harold asked, "*How can I write?*" (turns 135 and 151) indicating that he is the writer and he is in the process of writing down the list. At both times, there is a grammatical LRE associated with his asking of how to write down their idea. At turn 135 his use of '*made*' prompts several corrections to '*make*' from the others, which was not resolved until Harold finally said '*should make*' at turn 172. At turn 151, his question prompted Sunny to suggest '*culturally*' (turn 152) and then change the form to '*cultural*' (turn 154).

Grammatical LREs were not observed at all in listing tasks that involved listing only a word or expression, so it appears that coordinating sentence-level writing with speaking promoted attention to both meaning and form.

To sum up, the answer to the question of which tasks promoted students' attention to both grammar and vocabulary during interaction is tasks that involved student collaboration to write sentences. However, there were very few LREg's in the creative role play task or the simulation, with almost all LREs being vocabulary oriented. This suggests that the length and creative pressure of these tasks was too great for the time provided. The increased L1 during the planning stage of the creative role play and simulation tasks

suggests that students did not push to polish their language, just to finish. Values clarification and ‘longer-item’ lists produced a balance of meaning and form-related talk.

3.4.3 L1 use

One problem with using tasks in classrooms in which students share an L1 is that it will be used at the expense of L2 thereby reducing L2 learning opportunities (Carless, 2004, 2008). In this study, L1 (Korean) was not proscribed by the instructor, and it was observed in all task types. Often L1 was used to mediate vocabulary LREs as shown in Extract 3.1, turns 22, 25 and 26. In addition, while Harold was discussing grammar with Crystal, he used the Korean word ‘*myeong*’ (명) which means ‘*noun*’ in English (Extract 3.6, turn 161). He immediately supplied the English word but this was a rare case of an L1-mediated grammatical LRE.

More problematic for TBLT was that aside from resourcing language through LREs, L1 was often employed to convey task-necessary meaning, even though all of the task types aimed to promote L2 use through the expression of meaning. In other words, students used L1 to do what the task designer had wanted students to use L2 for. This was particularly prominent in the task planning stages of the creative role play and the creative simulation. For example, an examination of Extract 3.1 shows only 12 turns of L2-exclusive talk, two in interaction with the instructor. Section 3.4.1 reported that students used L1 to add new material to their presentation instead of practicing it.

These two tasks in particular involved public performances that required creativity and that aimed to entertain their classmates. Carless noted that tasks that were absorbing carried “a greater risk of student use of MT [mother tongue]” (Carless, 2008, p. 335) and this appears to be what happened in these creative tasks. Students used L1 to generate ideas and convince their team members of an approach, and only later, sometimes not until the final report (as in Extract 3.1) did they use English. Thus negating the benefits of task repetition had they rehearsed in English (see Bygate & Samuda, 2005; Bygate, 2001; Lynch & Maclean, 2000, 2001).

Another problematic observation from my observation notes was that in the last performance of the creative simulation, I noted that there appeared to be more L1 use than in previous tasks. This may have been due to the topic: Create a facility to improve student health on campus, or due to the fact that the task type had been repeated three

times. If the latter were the case, it is in direct contradiction to Kim (2013). I made no observations about increased L1 in the other task types over the length of the semester, however, it bears investigating.

To sum up L1 use, the tasks in which students used the most L1 were observed to be the more creative tasks. The amount of L1 increased as the time to present approached, and the phase of the task cycle intended for planning was often consumed by students adding content to their performance and using L1 to do so. This is in line with Carless' (2008) point about using L1 for more interesting tasks (p.335). The tasks with the least L1 use were listing and values clarification tasks. L1 used in these tasks was almost exclusively to resource L2 expression. This does not mean to imply that these tasks were not engaging. As shown in Extracts 3.6 and 3.7, students engaged with the tasks, and in the case of values clarification, expressed opinions and negotiated a consensus, almost exclusively in L2. The task requirements of a written English list or a revised statement written in English without a heavy requirement for creativity seemed to promote L2 use across all phases of the task.

3.5 Summary of the exploratory study

The exploratory study was carried out with the intention of describing how tasks situated in existing classrooms and integrated into existing programmes of study (i.e. ecologically valid tasks) would be carried out by students and how any changes from the planned task to the performed task might influence task performance. It also aimed to find tasks that used predominantly L2 and pushed learners to attend to both meaning and form. Finally, it introduced some questions that need resolving.

First of all, this exploration found that the three stages of Willis' (1996) task cycle, activity, planning and report, were not implemented by the students as planned and instructed by the teacher. More importantly, the planning phase never occurred as envisioned by Willis. The planning phase was expected to be a rehearsal for students' public presentations, yet the only way to get students to develop and improve the linguistic accuracy of existing content during the planning phase was for the teacher to explicitly tell a team not to add more content, and to stand nearby while they started to

rehearse. L1 was used extensively in the planning phase and this led to spontaneous, unrehearsed public reports. Other research (e.g. Ewald, 2004) suggests that making learners aware of the language learning value of tasks will change student behaviour. As failure to use planning time the way it was intended is a kind of behaviour, the next phase of this research will investigate how to encourage students to act during task-based interaction in ways that promote language learning.

This exploration revealed that having students collaborate to write a short final product (about one sentence in length) promoted a greater amount of focus on form than with other tasks. The values clarification task in particular provided a blend of meaningful talk that focused on the expression of opinions (opinion-gap) and on grammar-related talk. It was the one task that did both and as such was selected for further study in the next phase of this study.

It was also observed in the jigsaw tasks that students needed and appreciated having scripts of the input they had listened to previously. Because Willis (1996) urged consciousness-raising activities after the main task as part of a language focus, and students in this exploratory study were observed concentrating on the input scripts, the use of input was selected for further study in the next phase.

Finally, it was noted that L1 use appeared to increase in the final creative role-play task. It has been reported that task performance improves with repetitions of a task (Kim, 2013; Lynch & Maclean, 2000, 2001; White, 2011) yet most instructors would consider increasing L1 use as decreasing performance. As no study on task repetition has focused on ecologically valid, situated tasks (i.e. tasks designed for the learners doing them as part of their regular instruction and with their regular teacher and classmates) task repetition was selected as another item for study in the next phase of the research.

To sum up, the exploratory phase did what it was intended to do. It examined the ecology of classroom task performance and showed how students perform tasks designed for their situation and needs, and it reported findings that bear further inquiry. The next chapter outlines the methods of this further inquiry and the following four chapters detail the results.

CHAPTER FOUR: MAIN STUDY METHODS

4.1 Overview

The last chapter described the method and results of the first phase of the data collection. The chapter reported on six ecologically valid tasks and on how participants performed them. Similar to Carless' finding that when a task was more 'absorbing' there was greater L1 use (Carless, 2008, p. 335), when the tasks in the exploratory study had a competitive post-task report, more L1 was used in the task and planning stages. The exploratory study also found that processing input caused problems during jigsaw tasks and suggested providing scripts of the input for later analysis. Finally, it identified one task, a modified version of the values clarification (VC) task, which drew learners' attention to meaning while at the same time requiring them to attend to grammar during the task performance. When students did this task, there was limited use of L1. It was also easily integrated with the content of the curriculum. That is, the VC task provided a balanced learning opportunity for the students and it was therefore selected for further investigation.

The chapter starts first with a statement of the purpose of Phase 2 and a general indication of the approach. It outlines the nature of this investigation of the VC task. It starts by discussing the VC task in more detail and providing an example of one of the tasks used. It then discusses the participants. Then the specific materials used in the study are explained followed by an outline of the research design and implementation schedule. Data collection, coding and analysis procedures are at the end of this chapter.

4.1.1 Purpose statement of Phase 2

In this study, I investigated the pedagogical integration of the values clarification (VC) task into a pre-existing Sustained Content Language Teaching (SCLT) programme in a university in South Korea. The study set out to investigate task performance changes when using the VC task by taking a quasi-experimental approach in which three task-implementation conditions were varied across different pre-existing classes of students. The study investigated the VC task under the following implementation conditions:

1. The performance of a pre-task or a post-task input-processing activity (J. Willis,

1996b)

2. Task repetition: doing the same task on the same topic (Kim, 2013)
3. Procedural repetition: Doing the VC task on a new topic (Kim, 2013)
4. The presence of a metacognitive awareness (MA) raising task (cf. Ewald, 2004).

These implementation conditions were altered to investigate their effects on the following performance variables:

1. The use of targeted words and expressions from pre-task input
2. The number and type of Language Related Episodes (LREs)
3. Use of L1 during task performance
4. Accuracy of the formation a selected ‘task-natural’ phrasal verb (Loschky & Bley-Vroman, 1993), “*agree with*”, that was not explicitly targeted, but found naturally occurring in the input.

Although this study follows a quasi-experimental design, the use of pre-existing classes with their pre-existing curriculum and objectives, and the use of a site-validated task (Chapter 3) enabled students to experience the investigation as being a programme of study appropriate for this context. This claim was verified in the end-of-study interviews reported in later chapters in which students stated that the authentic input was a good idea (and offered ways to improve its use), that task repetition was a useful way to help their learning, and that the metacognitive awareness activity made them think about their task performance. This means that both Bronfenbrenner’s (1979, pp. 28–9) and Van Lier’s (2004, p. 195) criteria for ecological validity (see Section 2.2.3) have been satisfied because there is a match in the situational definitions between researcher (me) and participants (students).

4.1.2 Research Questions

The research questions addressed in the second phase of the study are listed below.

1. Does placement of an input processing activity (IPA) before or after the values clarification (VC) task affect task performance?

- a. Does the number of LREs change with IPA placement?
 - b. Does the number of turns in which L1 is used change with IPA placement?
 - c. Does the use of target items from the IPA activity change with IPA placement?
 - d. Does the target-like use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993) change with IPA placement?
 - e. Are the changes found in both original and repeated (delayed) task performances?
2. What is the impact of procedural repetition (Kim 2013) on VC task performance?
 - a. Does procedural repetition increase the number of LREs?
 - b. Does procedural repetition decrease the amount of L1 used?
 - c. Does procedural repetition increase the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?
3. What is the impact of task repetition (Kim 2013) on the task performance?
 - d. Does task repetition increase the number of LREs?
 - e. Does task repetition decrease the amount of L1 used?
 - f. Does task repetition increase the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?
4. What is the effect of the metacognitive awareness (MA) activity on VC task performance for groups who had the MA activity when compared to those who did not? (i.e. What are the across-groups effects of the MA activity?)
 - a. Does the MA task change the number of LREs?
 - b. Does the MA task change the amount of L1 used?
 - c. Does the MA task change the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?
5. What is the effect of the MA activity on task performance for groups who had the MA activity when compared to performances in earlier lessons? (i.e. What are the

within-groups effects of the MA activity?)

- a. Does the MA task change the number of LREs?
- b. Does the MA task change the amount of L1 used?
- c. Does the MA task change the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?

4.2 The Values Clarification (VC) Task

4.2.1 Describing the VC tasks used in the study

The values clarification task is a task designed to engage learners' knowledge of a topic and require them to talk to each other and reach a consensus on a series of controversial statements. As there are many possible consensuses the task is relatively open-ended, but it is convergent in that the participants must ultimately agree on one final version. Brinton (2003) described the values clarification task in the following way:

Values clarification tasks are relatively common in communicatively-oriented classrooms ... As a student-centred discussion activity, it provides a forum for students to voice their personal opinions about controversial topics. Because students are then asked to reach consensus in their groups, it also provides a forum for students to debate and persuade others of their opinions. Finally, it provides an opportunity for students to rationalize their opinions. (Brinton, 2003, p. 219)

Figure 4.1 is an example of a values clarification task used in the study (the full VC sheet with the full written instructions is in Appendix 1).

Figure 4.1. Sample Values Clarification Task, used in Lesson 3

Statement	SD	D	A	SA
(1) It is not a good idea to eliminate gender-biased language because language is too culturally important to change. Reason:	—	—	—	—
(2) Stereotypes about women lead to discrimination against women. Reason:	—	—	—	—
(3) It is more important to have good role models of powerful women on TV and in the real world than it is to change linguistic expressions. Reason:	—	—	—	—
(4) There have been very few real improvements for women, especially working women, over the past twenty years. Reason:	—	—	—	—
(5) There is less gender-bias in Korean language than in English. Reason:	—	—	—	—

4.2.2 Doing the VC task in this study

To do the task, the instructor passes out a single sheet of paper with a set of prompts, i.e. debatable statements, on it, and then gives the class time to read them. Students then indicate their level of agreement (strongly disagree (SD), disagree (D), agree (A) or strongly agree (SA)), by making a check (✓) in the blanks in the right-hand column. Students may then write a reason in the space provided if they wish. This is done silently. When the teacher tells the class to begin, the students talk to the other members of their team about each of the prompts. They discuss what each prompt means and try to reach a consensus about the idea expressed in the prompt. They are then expected to rewrite each prompt in such a way that it is something they can all strongly agree with. The revised sentences are written on a separate answer sheet and submitted to the instructor at the end of the VC task. The team then selects two prompts for presentation to the class. One member of the team reports their team's revised versions of these two prompts and explains their reasons for revising them. The class can discuss and challenge the presenter's opinion.

In the piloting for this study, it was observed that the last teams to present often continued talking and working on their final versions while the earlier teams were presenting, and earlier teams became bored and restless. To avoid this and to ensure that each team presented only what they had talked about, the presentations were given simultaneously. Each team's speaker stood up, facing only their team, and spoke directly into an audio recorder. Their team members were then asked to give the speaker feedback, and the presentation was then redone. The participants knew their instructor would listen to these reports as part of their classwork and feedback on these reports was given as part of the end-of-study debriefing session.

To sum up, the VC task involved: individuals reading each prompt silently and deciding on their individual opinion (pre-task); then, when talking was permitted, comparing and contrasting their opinions, persuading others, collaborating on a revised prompt and writing it down (task); finally, when the instructor ended the task, selecting a speaker, choosing two prompts to present (planning), and presenting it to the instructor via the recorder (report). This is expressed diagrammatically in Table 4.1, below.

Table 4.1. The Values Clarification task-as-workplan

Step	Stage (J.Willis, 1996b)	Sub-stage (time)	Expected activity
1	Pre-task	Read (5 min)	Students read silently and select their opinion by ticking their sheet They note down some reasons for their opinion on their sheet as well They may ask the instructor about the prompts or consult a dictionary
2	Task	Discuss (3min)	Students select a prompt to talk about Students express their level of agreement or disagreement with the prompt and give their reasons for their opinion Students compare their ideas and try to find mutual ground by suggesting changes to the prompt
		Write (1min)	Students come to an agreement One student writes down the final version of their revised prompt
		Move on	Students will select a new prompt and repeat the process
3	Planning	Discussion (3min)	Students will talk about which prompts to include in their report and what to say The presenter will practice
4	Report	First version (1min)	Presenter stands in front of their own group and speaks, on behalf of their group, for one minute into the voice recorder Group members listen and take some notes
	More Planning	Revision (3min)	Group members suggest changes Presenter notes them down and practices
	Report	Final Version (1min)	Presenter stands again and speaks for one minute into the voice recorder Instructor will listen to and mark this version

In general, a VC task may be done at the start of a textbook unit, to engage students' existing knowledge and prepare them for the theme of a unit (e.g. Brinton, 2003, p.219) or at the end of a unit to extend and consolidate knowledge of the topic (e.g. Sanabria & Sanabria, 2008, p.119). In this study the VC task was placed at the end of each unit. This ensured that all participants had at least a basic understanding of the topic.

4.3 The Participants

4.3.1 The Participating Students

The participants were first year undergraduate students enrolled in the second semester of an English-for-non-English-majors programme called Practical English 2. As with the exploratory study (phase 1), it was a SCLT programme and used the text *Academic*

Listening Encounters: Life in Society (Sanabria, 2004) as the primary text. The text provided the topical content, skills development exercises and vocabulary related to the topics. It was to be supplemented as needed by the instructors.

The participants were first-year undergraduate students who had enrolled in one of four classes (named classes A, B, C and D). Classes met twice per week, once for one hour (the first lesson of the week) and once for two hours (the second lesson of the week). Classes A and B met Monday afternoon and Wednesday afternoon. Classes C and D met Monday evening and Friday morning. As there may be an effect of time between lessons, classes A and D were grouped into one condition, hereby called 'Group 1', while classes B and C were combined into 'Group 2'. This balanced the number of participants (N=22 and N= 24) across conditions. For the target tasks, the classes were divided into small groups of three or four students creating seven groups in each condition. There were seven teams of participating students in each group.

4.3.2 Non-participating Students

If a participant opted not to take part in the study, he or she remained in the class. None of their speech was transcribed, counted or used in any way for this study. The speech of their group members was used but when they spoke 'XXX' was typed into the script. No LREs, L1 or any other variable in which they were involved were considered for analysis. Only one student opted not to participate.

4.3.3 The instructor and assisting colleagues

The researcher was the instructor for all the classes in this study. He had over 11 years' teaching experience at this university (see Chapter 1). Three colleagues from the Practical English programme assisted in making the input, in reviewing the VC prompts and Language Awareness activities, and with piloting.

4.4 Materials

4.4.1 Tasks

The values clarification task was described in Section 4.2. This section discusses the

making of the sets of prompts for each VC task.

Two sets of prompts for each chapter of the textbook were made. One set ('original') for the first task performance, and the second set ('repeated') for the repeated performance. The first set included five prompts in the form of debatable propositions that pertained to the topical content of that chapter. These prompts were checked using lextutor.ca (Cobb, n.d.) to ensure a similar number of words and a similar lexical difficulty across tasks. Then a second set of four prompts was made for each chapter by changing the structure or vocabulary of the prompts in the first set, thereby making a topically related set that was not identical. This was done to ensure that when students repeated a task they still had to negotiate about the language of the prompt and not simply repeat verbatim their answer from the prior lesson. To show the relationship between the two sets of prompts, three prompts used in the study are given below in both their original (1a, 2a and 3a) and repeated (1b, 2b and 3b) forms.

(1a) It is not a good idea to eliminate gender-biased language because language is too culturally important to change.

(1b) Eliminating gender-biased language will help change the culture in a positive way.

(2a) Stereotypes about women lead to discrimination against women.

(2b) Discrimination against women is caused by negative stereotypes about women.

(3a) It is more important to have good role models of powerful women on TV and in the real world than it is to change linguistic expressions.

(3b) It is more important to have good role models of powerful women on TV and in the real world than it is to change linguistic expressions.

Statements 1a and 1b are opposite in meaning, and a participant who agreed with 1a should disagree with 1b. However, 1b has additional information as well that will need discussion. 2a and 2b are similar and a participant should respond to both in a similar way. 3a and 3b are identical, and should be treated identically.

The prompts were shown to three other instructors in the programme who provided feedback on their topical suitability with respect to the textbook and on the wording of the

prompts. Once the prompts' content and wording were finalised, VC task sheets were made for each set (see Appendix 1 for an example).

4.4.2 Input-processing materials

The input referred to in this study consisted of an audio recording and accompanying transcript of native speakers (NS) performing the VC task. This recording and transcript were followed by comprehension questions to ensure the participants understood and a set of Language Awareness (LA) activities to focus the participants' attention to the NS's language. (A full set of input for one lesson is in the Appendix 2.)

To make the input, two other instructors and I performed a values clarification task on two of the items from each set of prompts for each chapter. I then transcribed the recordings to assess whether the task-as-workplan had been carried out, and the recordings were of similar length and level. Because some recordings were longer than others I decided to re-record the instructors performing the task, this time setting a time limit of 5:00 to discuss and finish one prompt to ensure a consistent length.

Then general comprehension questions about each script were made (e.g. "*Which of the three speakers agreed with the original statement?*"). After that, the transcripts of the instructors' talk were analysed for items that we felt would be problematic for students in that programme. Language Awareness (LA) questions were made from these following an inductive, consciousness raising (CR) approach (D. Willis & Willis, 1996). There were five general types of LA questions:

- Provide an equivalent expression to an item in the script
(e.g.: 4. Turn 2: *Gunther says, "is your intent to say". Give an equivalent expression.*);
- Find a synonym in the indicated part of the script
(e.g.: 4. Turn 34: *What word in this turn means 'suddenly'?*);
- Ask what an item in the script refers to from elsewhere in the script
(e.g.: 6. In turn 43: *What is meant by 'this type'?*)
- Scan the whole script for classes of items
(e.g.: 11. *Which phrasal verb (2-part verb) is most common in this script?*)
- Identify the implications of what the speakers were saying
(e.g. 7. In turn 19, *Shaun says, "sneak their opinions in". What does he imply by using the words 'sneak in'?*)

These questions were designed to focus students' attention on the words and expressions

in the input text that may be useful for the task. Students were not explicitly taught grammar rules and were not explicitly told to use these items. Because they were expressions used by native speakers, as they did the task, they were considered ‘task natural’ but not ‘task essential’ (Loschky & Bley-Vroman, 1993). The full list of target items can be found in Appendix 10.

Piloting of the finalised materials was carried out in another instructor’s class. At this time, I noticed that students generally finished the VC task in less than 20 minutes, and it took between 35 and 40 minutes to do the input-processing activity.

Training of the learners took place across two weeks. Students were put into teams of four or five and they discussed two prompts. They then presented their opinions and reasons to the class. A second training session was done the next week using two prompts on a different topic. This training was done to familiarise participants with interactive tasks and to get them used to talking to each other because many were strangers and came from language learning situations where they had little chance to interact (see Chapter 1).

In summary, for each chapter of the textbook there was a corresponding set of materials consisting of: (1) an original VC task with five prompts; (2) the input recording; (3) the input transcript; (4) the input LA activities; and, (5) a second set of four VC prompts for use in the repeated task.

4.4.3 Metacognitive awareness (MA) intervention materials

The metacognitive awareness (MA) intervention was a survey task designed to engage participants’ in critical reflection and examination of the VC tasks and on their own performance.

The survey used in the task is reproduced in Figure 4.2 below (Appendix 3 has the form with the instructions). Questions 1, 2, and 3 referred to the participants’ understanding of the instructor’s purpose behind doing the VC task in this way. Questions 4, 5 and 10 referred to specific things they felt or did when learning English. Question 6 asked about their approach to unknown language during task performance. Question 7 asked about their use of avoidance strategies. Question 8 asked about their L1 use. Finally, Question 9 asked about their beliefs about language and language learning. Talking about these things was expected to change performance in the VC task immediately after the MA task was completed (cf. Ewald, 2004).

Figure 4.2. The Metacognitive Awareness survey task questions

	Question	NO ← → YES
1.	Do you know why the teacher asks us to give our opinions about the topics from the book? If not, can you try to guess, please?	1 2 3 4
	Explain:	
2.	Do you know why the teacher asks us to try to (strongly) agree on one, new opinion? If not, can you try to guess, please?	1 2 3 4
	Explain	
3.	Do you know why the teacher makes us talk in English so much in class? If not, can you try to guess, please?	1 2 3 4
	Explain.	
4.	When you read the scripts and do the language awareness activities, do you try to remember new expressions, grammar or words?	1 2 3 4
	Explain why or why not.	
5.	When you hear your partners use an expression, word or grammar structure you don't know, do you try to use it yourself?	1 2 3 4
	Explain what you do when you hear or read some new words or grammar.	
6.	When you are talking in your group in class, are there ever times when you don't know how to say something in English?	1 2 3 4
	How do you feel when this happens?	
	What do you do when this happens?	
	What do you think the teacher would want you to do when this happens?	
7.	Do you sometimes just <u>agree</u> with your partners, even when you really disagree, just so the group can finish the task?	1 2 3 4
	Why do you feel you have to do this?	
8.	Do you sometimes use Korean when you are talking in class?	1 2 3 4
	What do you often talk about when you use Korean (the topic, English you don't know, something off-topic, etc.)?	
9.	Do you think talking a <u>lot</u> in English helps you learn <u>new</u> English?	1 2 3 4
	Explain why or why not.	
10.	What do you do to learn new English grammar, words or expressions while talking in English during class?	

4.5 Task implementation conditions (independent variables)

4.5.1 Pre-task vs. post-task input processing activity (IPA)

One of the aims of the research was to investigate the role of input and whether placing it before or after the main task had any impact on task performance. To address this, Group 1 was assigned to the task-then-input condition and Group 2 to the input-then-task condition for the first chapter of the study. The classes' treatment conditions were switched at the end of the next chapter of their textbook, giving the study of input

placement a within-subjects repeated measures design.

The top portion of Figure 4.3, labelled ‘Original Task’, shows the procedures for each of the conditions within a lesson. The 11 steps took approximately 60 minutes to complete.

Figure 4.3. Comparison of activities in input-then-task and task-then-input conditions

First Lesson: Original Task	
Input-then-task	Task-then-input
Pretask 1. Read the original prompts silently 2. Decide on own opinion and reasons Input 3. Listen to and read input 4. Reread input at own speed (silently) 5. Answer comprehension questions (CQ) 6. Do Language Awareness (LA) activities 7. Teacher-led plenary of answers to CQ and LA VC task 8. Discuss prompts and revise the prompts Report 9. Choose a presenter 10. Present, get feedback from group members 11. Present final version	Pretask 1. Read the original prompts silently 2. Decide on own opinion and reasons VC task 3. Discuss prompts and revise them Report 4. Choose a presenter 5. Present, get feedback from group members 6. Present final version Input 7. Listen to and read input 8. Reread input at own speed (silently) 9. Answer comprehension questions (CQ) 10. Do Language Awareness (LA) activities 11. Teacher-led plenary of answers to CQ and LA
Subsequent Lesson: Repeated Task	
Pretask 12. Read the repeated prompts silently 13. Decide on own opinion and reasons VC task 14. Discuss prompts and revise them Report 15. Choose a presenter 16. Present, get feedback from group members 17. Present final version	Pretask 12. Read the repeated prompts silently 13. Decide on own opinion and reasons VC task 14. Discuss prompts and revise them Report 15. Choose a presenter 16. Present, get feedback from group members 17. Present final version

The design permitted the investigation of difference between the two groups’ task performances on the basis that any differences could be attributed to doing the task using the input as a resource for topical content and language (input-then-task) or doing the task with no additional topic knowledge or linguistic examples (task-then-input).

4.5.2 Task repetition (the original task vs. the repeated task)

At the start of the following lesson after the VC task was done, under either input-then-task vs. task-then-input conditions, the VC task was done again, using the second set of prompts. This second set was called the ‘repeated’ task. As there were four prompts in the repeated task and not five, less time (10-12 minutes) was allotted for completion.

Note that the input was provided to both groups in the original lesson, and learners were not reminded of the LA items in the repeated condition. Thus, if performance in the repeated task differed between groups, the differences could be attributed to the *order* of input placement in the previous lesson.

4.5.3 Metacognitive Awareness (MA) intervention

In the fifth lesson of the study, the metacognitive awareness task was carried out by Group 1. Learners in this group performed a survey activity in which they asked each other about their approaches to the values clarification (VC) task (see Figure 4.2 and Appendix 3). This survey implicitly raised the participants' awareness of the main purpose of the task – to learn language. The other half of the participants did another activity from their course textbook, and took the same survey after the end of the study.

For this survey, participants were placed in different teams from their values clarification (VC) teams and were given the survey. Participants silently read the instructions (found in Appendix 3) and circled their agreement level, 'No \leftrightarrow Yes'. They then could, if they wished, jot down an explanation. After about ten minutes, students were asked to talk to each other about each of the items, share their ideas and concerns. Then a representative of each group presented a summary of their ideas, followed by a general discussion.

4.6 Overall research design

The three implementation conditions (input, repetition and metacognition) were combined to give the research design schedule shown in Figure 4.5. Piloting was done in the 3rd, 4th and 5th weeks of the semester because national holidays interrupted some classes' lessons. Lessons 1~4 took place from the 6th-8th week of the semester. They were followed by midterm examinations and group discussion tests. The MA+ task was done in the 11th week.

As this was classroom research, all lessons were scheduled by the university authorities and there was no opportunity to change times and dates. This led to the issue of classes A and B meeting on Monday and Wednesday, and Classes C and D meeting on Monday and Friday. To try to balance this time difference, Classes A and D were put into the same

treatment conditions and Classes B and C were put in the other.

Figure 4.4. Overview of lesson, chapter and lesson condition for the study

Lessons	Task Repetition Condition	Input placement condition			
		Class A	Class B	Class C	Class D
Piloting 1	Original	Task-then-input			
Piloting 2	Repeated				
Lesson 1	Original	Task-then-input	Input-then-task	Input-then-task	Task-then-input
Lesson 2	Repeated				
Lesson 3	Original	Input-then-task	Task-then-input	Task-then-input	Input-then-task
Lesson 4	Repeated				
Lesson 5	Original	(MA+) Task-then-input	(MA-) Input-then-task	(MA-) Task-then-input	(MA+) Input-then-task
Lesson 6	Repeated				

Notes: MA+ = metacognitive awareness intervention done, MA- = metacognitive awareness not done

4.7 Data collection and analysis procedures

Data collection started from the second week of the semester, after students had been briefed about the research, given information sheets and signed consent forms. This was done to familiarise them with the presence of the recorders and video camera.

The principal source of data for the study was classroom recordings taken during every lesson. As in Phase 1, above, small voice recorders were placed at every team's desk for the entire lesson, allowing the researcher to not only listen to the VC task performances, but also student talk during the input plus language awareness (LA) phase, the MA task and other classroom activities.

One video camera was set up near the instructor's station at the front of the room. It was focused on the team nearest the instructor to record the actions, gestures and talk of one team each lesson. The microphone on the camera was not sensitive enough to catch the softer talk in the videotaped teams, but the video could be synchronised with the audio

recording when needed using ELAN (ELAN, 2013; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

While teaching, the instructor kept notes about classroom seating arrangements, main timings of events and his thoughts as each lesson progressed.

Students also provided written data in the form of their final, revised versions of each VC prompt, which were scanned prior to being marked and returned. Likewise the MA+ task survey sheets were also collected.

At the end of the semester, the instructor held a debriefing session in which the participants were reminded of the purpose of the study, reported preliminary findings, and provided feedback on their performance. In this session, they also listened to sample recordings of their task performances. Participants were encouraged to ask questions to the instructor and to comment on their own development. These sessions were not recorded.

After the study was completed, students were asked to volunteer for end-of-study stimulated recall interviews in which portions of their talk were played for them and they answered questions about their task performance (see Appendix 4, below). These were semi-structured interviews and the participants were invited to comment, ask questions or make suggestions.

A potential weakness of this design is the lack of a pre-test, post-test and delayed post-test for general proficiency gains or for target items. I considered doing this, but felt such tests would interfere with the study. First, ethically I had to ensure that all students did the same work in the course. Therefore, the design was counterbalanced, so ultimately all students did all conditions. This meant that overall proficiency gains should be equal across groups. Second, testing target items before and after each VC task would focus students' attention to the items more than just doing the task would. Any language gains or behavioural changes could not be attributed to the task implementation condition only, but to the presence of a test. (Indeed, students recommended giving a test to help them learn more items, see Chapter 6). Moreover, an initial pre-test and a post-test at the conclusion of the study would be biased toward the items encountered later in the study. In this way, the ecological setting precluded a more robust cause-effect study of the effects of implementation condition on *learning*, and limits the study to the effect of

implementation condition on task *performance* with the assumption that improved task performance will enhance learning outcomes (more on this in the limitations, Sect. 9.3, below).

4.8 Dependent variables

After each lesson, the video recording of the class was viewed and timings of classroom events were recorded, seating and movement charts were completed if necessary, and the researcher compared his observation notes with the video. Then, individual recordings were transferred from the recorders to a computer for transcription. Transcripts were made using voice walker (Du Bois et al., 1998).

First, all of the in-task talk for randomly-selected teams, one from each class, was transcribed for each lesson throughout the semester. These were then given a preliminary analysis for salient features of the talk.

Once this had been done, all the recordings were analysed by listening to them, identifying the salient features, transcribing the relevant part of the script and classifying the features according to the criteria below. The features of in-task talk examined were: LREs (grammatical, lexical, phonological), use of targeted forms from the LA activity, L1 use, and the correct formation of a task-natural phrasal verb.

4.8.1 Language Related Episodes (LREs)

As discussed in Chapter 2, LREs are points in the interaction when learners “talk about the language they are producing, question their language use, or correct themselves or others” (Swain & Lapkin, 1998, p. 326). LREs were identified in the learners’ in-task talk and classified as to their focus: on grammar, vocabulary or pronunciation.

(1) Grammar-focused LRE (LREg)

Decisions about whether the LRE was focused on grammar or vocabulary were often difficult. A grammatical LRE (LREg) was coded if the item of difficulty and its resolved form were related to the morphosyntax of the utterance. Extract 4.1, shows part of a discussion of the prompt “*Mass media is not as good for us as the lecturer claims.*” In

turn 23, Mary dropped the subject (mass media), which they all know, and has supplied a possible ending to their revised sentence. Monty suggested a different ending to the sentence in turn 24. Eunice then prompted the LREg by nominating the incorrect *have* for *has* in turn 25, which Monty then rejected it and resupplied the correct form. Eunice repeated the correct form, indicating she has heard.

Extract 4.1. LREg, selecting *have* or *has*.

Turn	Speaker	Utterance
23	Mary	is bad for us
24	Monty	uh has bad things
25	Eunice	ha- have bad
26	Monty	HAS bad
27	Eunice	has bad things

(2) Vocabulary-focused LRE (LREv)

An LREv is an episode that focuses on the meaning of a word or expression in the utterance. Extract 4.2, below shows Eunice helping Monty with the word *individual*. Monty could not come up with the word in turn 7 and said “*im-imvis ?*”. The rising intonation indicates his uncertainty with the term. Eunice suggested *individual* in turn 8. Monty responded, ‘*ah*’ to express agreement and repeated the term to reinforce it to himself.

Extract 4.2. LREv, nominating the word *individual*.

Turn	Spkr	Utterance
6	Eunice	because her points were good . it’s true I- I h* agree with her ideas
7	Monty	I think this depends on . uh im- imvis ? gain
8	Eunice	individual ?
9	Monty	ah individual

Often LREv involved the use of L1. Extract 4.3 shows Sunny using the word, *disregard*, which Winnie did not know. She asked in turn 82, but Sunny thought Winnie had misheard her. Winnie clarified her question in turn 84, and Sunny gave a translation in turn 85. In turn 86 Winnie acknowledged the translation but never repeated or used the term *disregard* herself.

Extract 4.3. An LREv mediated with L1 (‘*disregard*’)

Turn	Speaker	Utterance
81	Sunny	boys boys tend to: disregard woman teachers

82	Winnie	wha- wha- what tend to ?
83	Sunny	tend to disregard
84	Winnie	disregard is what ?
85	Sunny	<i>LImusihada</i> 𐌆𐌵𐌶𐌵𐌹𐌸𐌵𐌹𐌸𐌵𐌹𐌸𐌵𐌹𐌸 (tr= disregard)
86	Winnie	ah:: ah but

(3) Pronunciation-related LRE (LREp)

An LREp focuses solely on the correct pronunciation of an item. Extract 4.4 is an example of an LREp. Eunice repeated the word ‘wisdom’ twice. The first was correctly pronounced and in the second she altered the pronunciation and indicated uncertainty with it through rising intonation. Monty then provided the correct form, which Eunice repeated.

Extract 4.4. LREp about the pronunciation of ‘wisdom’

Turn	Spkr	Utterance
24	Eunice	but I think they have many information and knowledge h* uh than PAST but they don’t know how to use them and h* they are not wise . only knowledge ? not wisdom (/wɪz dəm/)- wisdom (/waɪz dəm/) ?
25	Monty	wisdom (/wɪz dəm/)
26	Eunice	wisdom (/wɪz dəm/)

4.8.2 L1 use: on and off topic

L1 use has been identified as a major issue in the research on the implementation of tasks in classrooms, particularly monolingual classes (Butler, 2011; Carless, 2004, 2007, 2008, 2009). Willis (J. Willis & Willis, 2007, pp. 26, 220–1) accepts that L1 will be used during the task. She advises that the post-task report needs to be done in English. Pinter (2005) had young children do the task in L1 first and then change partners and do it again in English. However, for tasks such as the values clarification (VC) task, in which development of the capacity to do the process of reaching a consensus in English is the rationale for doing the task, overuse of L1 may reduce the learning opportunities afforded by the task. So, the amount of L1 used by students was tracked to determine whether the implementation conditions (input, repetition and metacognitive awareness) increased or decreased its use.

A broad quantification of L1 use was used. Any turn that included L1 (excluding LREs,

proper nouns, and interjections) was counted as an L1-turn³. L1-turns were later sub-classified into on-topic and off-topic L1. The number of turns using L1 per task was calculated. As the amount of time given for each VC task was not always constant across classes due to classroom factors (although it was kept close) and the repeated lessons had fewer prompts and took less time, the number of turns of L1 was standardised to a ratio per 10 minutes of talk.

I acknowledge the limitations to this approach. Particularly that it is very broad in and does not give precise detail about the length, topic or function of the L1 used. However, L1 use represents a choice for these learners and the number of turns using it is essentially the number of times they chose it over attempting to use the target language. In this respect, number of turns of L1 (per 10 min) provides an indication of behavioural choices.

4.8.3 Use of a specified task-natural lexico-grammatical structure: *agree with*

The decision to track the incidental use of the phrasal verb '*agree with*' was based on students' performance observed in the exploratory study (Chapter 3). Students often made mistakes using the verbs '*agree*' and '*disagree*'. There were many instances of BE+AGREE, such as "*I'm agree ...*", and AGREE+TO+NP, such as "*I agree to your opinion.*" Phrasal verbs are known to be challenging for learners (Nassaji & Jun, 2010), so if the use of this one phrasal verb improved through exposure to task input or through task repetition or both, this would be evidence of implicit learning from task performance.

The correct form "AGREE+WITH+NP" appeared in the task instructions, in the native speakers' scripts used for input and the in the comprehension questions about the input no less than six times with every set of input. This happened spontaneously during the native speakers' task performance and naturally in the written instructions and questions. There was no deliberate seeding of the input, nor was it made salient in any way in the script, the discussion or from instructor feedback (until the debriefing session). Because of this, *agree with* was considered to be a task-natural form, helpful, but not essential, to the VC task (Loschky & Bley-Vroman, 1993).

³ An exception to this was made for one prompt which read, "*There is less gender bias in Korean language than in English.*" Examples of Korean words or expressions that were integrated into the L2 talk were not counted as L1 turns.

The total number of times that *agree* appeared with the correct preposition as a percentage of the total number of times *agree* appeared with any preposition or when a preposition was necessary and not supplied was calculated. When *agree* was used correctly with no preposition and none was required, it was not included in the data.

4.9 Reliability

To ensure reliability, a discursive process was followed. Video of the class procedures was watched and compared to my in-class observations. Timings of all VC tasks were identified and group membership double-checked. Then all VC task talk was listened to and the main events were identified along with their timings. I repeated this process three times to ensure nothing was omitted. Then transcripts of all VC task talk activity along the entire semester were made for 17 tasks (20% of the dataset). These transcripts were then checked against the audio-analysis to ensure accuracy. After names of the participants and any identifying information about them from the scripts had been removed, the transcripts were sent to a second researcher who was familiar with the variables and with Korean learners of English. The second rater identified and coded LREs independently of the researcher. The two raters discussed all discrepancies until consensus was reached and the above rating definitions and policies were formed. I then analysed the remainder of the data. Other variables (e.g. L1 use) were not double-rated as deciding on an utterance's membership in one of these categories is unambiguous. Once variables had been coded and counted, a 10 per cent sample was reanalysed by the researcher to ensure intra-rater reliability ($r=0.91$). These final counts were then used for analysis.

4.10 Chapter Summary

This chapter outlined the methods used in the study. First a description of the values clarification (VC) task was given. This task was chosen due to the finding in the exploratory study (Chapter 3) that it balanced learners' focus on expression of meaning

with a focus on form more than other tasks did. It had also been found to be easily integrated into an existing programme. After describing the VC task, this chapter then described the participants and the development of the input and other materials. The implementation of the tasks into the study was described next, followed by a description of the data, analysis and coding and steps taken to ensure reliability. The next chapter presents a qualitative description and discussion of the VC task-in-process.

CHAPTER FIVE: VALUES CLARIFICATION TASK PERFORMANCE

5.1 Introduction

In this, the first of the four results and discussion chapters I analyse and evaluate the values clarification (VC) task as it was performed by the students in the study. It lays the foundation for the following results chapters that build on it. It starts by presenting two teams' performance of one prompt of a VC task in order to highlight their different approaches to the task and orientation to the task as a learning situation. After this, a more general analysis of the learning opportunities in the VC task is presented using examples from the same two teams along with others. Then the analysis is triangulated through the voices of the participants before conclusions are drawn. The aim of the chapter is to provide the reader who may not be familiar with the VC task, a clear understanding of the interactive processes that drive it and how this motivates language learning.

5.1.1 The data set

The intervention study took place across six lessons. Forty-six participants were divided into 14 teams, and formed into two groups based on existing class membership (Chapter 4). Each team had an audio recorder placed in the centre of their desk that recorded all the talk at that location for the lesson. This included the VC tasks and all other classroom activity. One team was also videotaped per lesson. In addition, ten students volunteered for end-of-study, semi-structured interviews. This totalled 158 hours of recordings. In addition, the participants' rewritten VC prompts were collected at the end of each lesson, as were the paper-based survey forms for the metacognitive awareness (MA) task (Chapter 8), and I kept observation notes of class events.

Chapter 4 details the data analysis process, but to remind the reader, an iterative process was followed. First, my observation notes were read and annotated, then the video was reviewed to crosscheck my observations, and then the audio recordings of team talk were listened to and transcribed. Finally, coding took place. A similar process was followed with the interview data.

5.2 Performance of the task

This section describes the performance of the VC task. I first describe the task-as-workplan and then give a brief descriptive overview of the way students performed the task, i.e. the task-in-process. Then I provide a more detailed account, drawing first on an example task performance (Team A5) chosen at random using www.random.org (Haahr, 1998) and then contrasting it with a performance of the same task done by another team (Team D1).

5.2.1 The values clarification task-as-workplan

The values clarification (VC) task-as-workplan can be characterised as an opinion-gap driven, convergent, restricted-open response, consensus task (Ellis, 2003; Samuda & Bygate, 2008). The plan was that to do the task, participants would first read and think about the prompts silently for a few minutes before starting to talk. Each participant could then bring a different opinion, creating an opinion gap to drive the interaction. Next, the participants would ask each other about their opinions about a prompt, express their own opinion and reasons for holding this opinion, and then work to resolve the opinion gap and reach a consensus on a new version of the prompt that they all strongly agreed with. They would then write the new, mutually agreeable version of the prompt on an answer sheet. In the sense that this new version of the prompt could use any words or expressions and take any point of view, it was open-ended, but because the task involved rewriting an existing prompt, the new version usually involved many of the same words as the original. In this way, the response is restricted – neither open nor closed. Finally, one member would present their views to the class in a one-minute public report.

5.2.2 The values clarification task-in-process

As found in the exploratory study (Chapter 3), in the second phase of the study, students typically performed the VC task in a manner consistent with the workplan (see Table 4.1). First, it was observed that all teams made revisions of the prompts that all their members strongly agreed with. No team ever produced no revisions, although some failed to revise all five prompts. Moreover, all members of all teams participated. There were no cases of a student remaining silent through the discussion of a single prompt, and there were no cases of a student only using L1 for an entire prompt or for the majority of their

talk for that prompt.

In their effort to produce an acceptably revised prompt, teams tended to approach the task in one of two ways (see below). The first approach was an opinion-oriented approach. When discussing a VC prompt in this way, each member of the team would express an opinion and engage with the ideas expressed by their partners prior to beginning to revise the prompt. The second approach was a revision-oriented approach. Here, they sometimes stated their opinion with little elaboration or reasoning, but, more often, they skipped expressing an opinion entirely and immediately began to revise the prompt. If they did state an opinion, they did not refer to the opinions their team members expressed, and just provided their own ideas for the others to hear. When revising the prompt students simply nominated words or phrases that were either accepted or rejected by the team. During this acceptance/rejection process, further details were provided by the speakers.

These two approaches to the VC task were not consistently followed by any team. There was considerable variation in the approach to the VC task both between different teams, and within the same team, as they worked on the five prompts of a VC task. Teams often took an opinion-oriented approach for some prompts and a revision-oriented approach for the remaining prompts. This variation appeared across the various task topics.

Therefore, although the VC task was consistently completed, teams often took different approaches to task completion, even within a set of prompts. These different approaches provide different task performances, based on different interaction patterns, leading to different learning opportunities each time the VC was performed. This implies that for theories of task performance, assuming a mechanistic ‘task-causes-output’ cause-and-effect relationship, irrespective of the students, may be misguided and a closer look at the process of performance in context will be needed.

5.2.2.1 Opinion-oriented approach: Team A5, Lesson 5 discussing prompt 1

This section describes the task performance of Team A5 as they do their task. This team’s talk for this prompt typifies an approach in which all members give their opinion prior to attempting to revise the prompt (an opinion-oriented approach).

As seen in Extract 5.1, this team’s successful completion of the VC task prompt involved them collectively understanding the meaning of each prompt, individually providing and comparing opinions, and then generating a consensus response. In the extract, the three

students: Lucky, Cathy and Jessie (Team A5), are starting their task performance in Lesson 5. They are discussing the prompt “*Mass media is not as bad for us as the lecturer claims. What she said was too one-sided.*” The ‘lecturer’ mentioned in the prompt referred directly to the course material they had studied in the previous lesson (Sanabria, 2004, pp. 91–96).

Extract 5.1. Team A5, Lesson 5, discussion of prompt 1

Turn	Speaker	Utterance
16	CATHY	let's start number one . I agree
17	LUCKY	why
18	CATHY	because uh
19	JESSIE	what she said was TOO one sided ? but it's okay because I don't know anything about it (hh) I didn't know so she saying is a new information to me so it will be the uh start point about knowing information
20	CATHY	(silently reads prompt) so then you said disagree ? no ? we have to say . (voice trails off)
21	JESSIE	mass is not bad- not as bad L1 <i>nappeuji anh-eunde</i> <i>나쁘지 않은데</i> (tr= not bad)
22	CATHY	but reason is a little bit different
23	JESSIE	really ?
24	CATHY	oh oh oh okay I agree
25	LUCKY	why?
26	CATHY	because uh last time I talk about to you ? but I didn't trust mass media TOO much like the lecturer claims h* but I think lecturer's claims so: bad too one sided . because h* uh mass media has a benefit . TOO I think . so I agree
27	LUCKY	okay . I strongly agree . because sometime mass media teaches us some useful information that we could use in our own lives . so but the lecturer only said about the negative way- negative side of the mass media but it has a positive side too so I strongly agree
28	CATHY	then change it .
29	JESSIE	lecturer's claim is what- what is exact meaning ?
30	CATHY	uh I don't remember exac-
31	LUCKY	-lecturer said the mass media teaches kids uh bad things and that- and also mass media makes people to go stupid
32	CATHY	oh really ?
33	LUCKY	YEAH mass media has a lot of violence
34	CATHY	ah ah //okay okay okay\\
35	LUCKY	//that's lecture's claim\\
36	CATHY	it's your turn JESSIE
37	JESSIE	WHAT ?
38	LUCKY	are you angry hh
39	JESSIE	I have to write ?
40	LUCKY	don't be angry
41	CATHY	yes
42	LUCKY	//because you have to\\ ₁
43	JESSIE	//I use left \\ ₁ hand
44	LUCKY	//okay\\ ₂
45	CATHY	//okay\\ ₂ mm we are AGREE bu:t h* hm but strong we have to change it . strongly agree

46	LUCKY	mm (to T) uh how can we change this ?
47	T	how about the 'too'
48	LUCKY	ah
49	CATHY	change this word ?
50	T	yeah
51	LUCKY	what she said was TOO one-sided
52	JESSIE	that's just true
53	T	then you would strongly agree ?
54	CATHY	ah: it's to strongly agree uh I- I- ah have SAME . same opinion with that lecturer claims h* so I think cause she said . uh little bit . one-sided or she talk about . too based on biased . //thinking\\
55	JESSIE	//uh how about\\ how about this . eliminate all what she said was too one-sided eliminate and just mass media is not as bad for us as the lecturer claims
56	LUCKY	uh okay
57	CATHY	okay simple (4.0) number two young people are becoming more and more stupid because they watch tv too much and don't know how to use the information they get

This extract demonstrates how the VC task was performed. The students discussed the prompt and rewrote it to a mutually agreeable form. Their interaction is comprised of four phases, each of which is discussed in detail below: understanding the prompt, expression of opinion, negotiation of a mutually agreeable revision, and task-organising talk.

Understanding the prompt (turns 16-24)

The first phase of this extract shows the team working to understand the prompt's language and the content. Here, the participants began by identifying the prompt and then starting to give their opinions about it. Cathy failed to fully express her opinion (turn 18) because she had not correctly understood the prompt. She and Jessie disagreed about the meaning of the prompt and Jessie used L1 to resolve the issue (turns 20-24). Cathy was then able to express her ideas fully (turn 26).

Within this sequence, Jessie expressed her opinion at turn 19. She casts the lecture recording from their textbook as new information, even though course attendance sheets and my notes confirmed that she had been present in the previous lesson. Later, Jessie indicated she still did not understand the referential meaning of the prompt (turn 29). Cathy admitted that she did not know it exactly either (turn 30). They did not remember what the lecturer had said. This was resolved from turn 29-33 by Lucky. This interaction demonstrates the relationship between the VC task and the course content, in that knowing the course material was also needed for successful VC task performance. (See section 5.3 and 5.4).

Expressing opinion (turns 26-28)

The second phase of the VC task was expressing opinions. In Extract 5.1, all three participants used extended turns to give their opinions: Jessie (turn 19), Cathy (turns 26 and 54), and Lucky (turn 27). These turns are approximately equal in length (approx. 40-50 words). While one member gave her opinion, the other two listened without interrupting. In the immediately following turns: Cathy (turn 20, 28), Lucky (turn 27) and Jessie (turn 55), the next speaker typically indicated engagement with the speaker's opinion. Cathy's move, in turn 28, was the only one of the four turns that immediately followed an opinion that did not contain a direct reference or acknowledgement of the preceding opinion. These lines indicate that all members used the VC task as an opportunity to express their opinion and engage with their partners' opinions.

Negotiating consensus (turns 45-56)

In the third phase of the task, students try to reach consensus on a rewritten version of the prompt. The rewriting of the prompt was an essential part of the VC task as it was anticipated that here participants would attend to the language needed to construct a revised version, and in so doing, engage in negotiation of meaning (Long, 1996) and LREs (Swain & Lapkin, 1995, 1998). Extract 5.1 revealed a different tactic. At turn 45, Cathy voiced her concern that although she agreed, her level of agreement was not strong enough to say 'strongly agree', indicating they needed to modify the prompt but not change it dramatically. Lucky asked the teacher who had walked past the team, and after a brief exchange, he left the participants to resolve it themselves. So, in turn 54, Cathy said that she had reservations about part of the prompt, she did not think it was *too one-sided*, just a "*little bit one-sided*". In turn 55, Jessie suggested deleting the entire last clause, and the other two agreed without further negotiation. Their final revision read, "*Mass media is not as bad for us as the lecturer claims*" (Figure 5.1). The negotiation of consensus in Extract 5.1 was not as extensive as predicted. The task-as-workplan anticipated there would be more engagement with the language of the prompt. In Extract 5.1, the only engagement with the prompt on a linguistic level during revision, was Cathy's rewording of 'too' to a 'little bit' one-sided, and neither term ended up in the final version (Fig. 5.1).

Figure 5.1. Team A5's final version of prompt 1

1. Mass media is not as bad for us as the lecturer claims.

Task organisation

In the fourth phase, task-organising talk, the students negotiated member roles and task activity. This phase was often discontinuous, as students engaged in task organisation at various points in the interaction. In Extract 5.1, member roles were negotiated using L2, although in other teams L1 or gesture was used. Turns 36-44 were a discussion of who would do the rewriting; with Cathy pointing out that it is Jessie's turn to be the writer. Jessie's threat that she would write left-handed, although humorous, was not acted on (as the video revealed). Moving across phases included starting a new prompt and moving to the revision stage. Starting a new prompt meant signalling which prompt they were talking about (at turn 16) or signalling the prompt and then reading it aloud (turn 57). In some cases the prompt was read aloud without signalling which one it was, but this did not happen in this extract.

Overall, Team A5 performed this prompt successfully, with most of their turns focussed on expressing and understanding each other's opinion. They assigned roles within the team and moved through the stages of the task in L2. However, there was limited discussion of the language in the prompt while they were revising. In the next sub-section, Team D1's performance shows a different focus, one toward collaboratively negotiating the final revised statement.

5.2.2.2 Revision-oriented approach: Team D1, Lesson 5, discussing prompt 1

This section describes the task performance of Team D1 as they do their task. This team's talk for this prompt typifies an approach in which task talk centres on the revision and not on their initially stated opinion. Although all three members give an opinion at the outset, there is no mutual engagement with their partners' ideas until they try to revise the prompt. In this sense, the drive for (written) consensus is pushing their talk in a way unlike that of Team A5, above.

Extract 5.2, below shows team D1. Monty (male), with Eunice and Mary (female) participants, discussing the same prompt as in Extract 5.1: "*Mass media is not as bad for*

us as the lecturer claims. What she said was too one-sided.”

Extract 5.2. Team D1, Lesson 5, discussing prompt 1

Turn	Spkr	utterance
1	EUNICE	what do you think about number one ? hh
2	MARY	I disagree
3	MONTY	I agree
4	EUNICE	I- I agree
5	MARY	(2.0) agree ? why ? why you disagree ?
6	EUNICE	because her points were good . it's true (MARY: mm) I- I (2.0) h* agree with her ideas
7	MONTY	I think this depends on . uh im- imvis gain
8	EUNICE	individual ?
9	MONTY	ah individual
10	MARY	I think she has already stereotype (MONTY: mm) about mass media so
11	MONTY	I think
12	MARY	so she's lecture is so: biased
13	MONTY	mm then
14	EUNICE	I think mass media is bad for us as the lecturer but I agree with that what she said was too one-sided she didn't mention about the advantages of mass media
15	MONTY	then: h* (whispering)
16	MARY	um: (2.0) mass- mass media is (1.0) also mass media have um bad things but she's lecture is so biased
17	MONTY	but ?
18	EUNICE	it is
19	MONTY	-but what she said was too one-sided
20	MARY	ah one-sided um so //what\\
21	MONTY	//so mass\\ media
22	MARY	mass media
23	MONTY	has uh
24	MARY	is bad for us
25	MONTY	uh has bad things
26	EUNICE	ha- have bad
27	MONTY	HAS bad
28	EUNICE	has BAD // things \l ₁
29	MONTY	//things\\ ₁
30	MARY	//things\\ ₁
31	MONTY	as the lecturer claims
32	MARY	uh mm
33	MONTY	(4.0) has SOME bad things some bad things
34	EUNICE	bad points
35	MONTY	uh uh bad things bad aspects
36	EUNICE	uh aspects okay
37	MONTY	s: //some bad aspects\\ ₂
38	EUNICE	//some bad aspects\\ ₂
39	MARY	//but she say was\\ ₃
40	EUNICE	//as the lecturer claims\\ ₃
41	MONTY	mm mm mm as the lecturer claims
42	MARY	but she said was too one-sided
43	MONTY	but WHAT she said what she said was TOO one-sided
44	MARY	(6.0) and number two

Understanding the prompt

Unlike Team A5 in Extract 5.1, Team D1 in Extract 5.2 displayed no misunderstandings due to the language or course content.

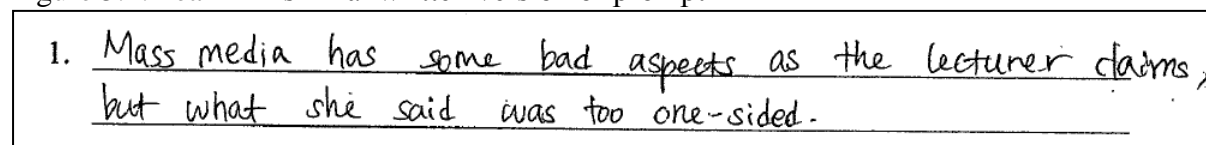
Expressing opinion (turns 6-16)

Again, in contrast to Team A5, Team D1 did not engage in extended expression of opinions, which were so common in Extract 5.1. The participants all give an opinion: Mary at turns 10, 12, and 16; Monty at turns 7 and 9; and Eunice at turns 6, and 14, but they did not engage with each other's opinion and moved quickly to the revision process. For example, Monty at turn 7 gave an opinion on the prompt without referring to or acknowledging Eunice's turn 6. A vocabulary LRE in turns 7-9, in which Eunice provided a word that Monty did not know, indicates that the members were listening to each other, but they did not build on each other's ideas. Of the turns that expressed opinion, Eunice, at turn 14 gave the longest explanation, 31 words, far shorter than those in Extract 5.1.

Negotiating consensus (turns 21-43)

Team A5 (Extract 5.1) showed very little negotiation of consensus, whereas Team D1's (Extract 5.2) negotiation consisted of half the turns. They started to revise the statement at turn 21, when Monty is speaking. The video revealed that Mary simultaneously indicated to Eunice that she should write down what Monty was saying (Fig. 5.3). Mary's gesture indicated they had started to revise. Over the next 22 turns, they collaborate through a series of suggestions and counter-suggestions for words until they arrive at the final version (Fig. 5.2). Many of their turns were sub-clausal, overlapped with other speakers, and latched directly onto the language of the preceding turn, creating an extended co-construction (Ohta, 2001). For example, in turns 33-36, the participants nominate '*things*' and '*points*' before agreeing on '*aspects*'. These turns contained no evaluative comments, no clear rejection of the preceding word and did not indicate any lack of comprehension. The participants were focused on the language of their final answer through what might be characterised as a 'brainstorming' approach to vocabulary; they nominated items aloud until they found the best one. The entire process of rewriting the prompt took place in this way – a series of shorter collaborations with language related episodes (LREs) mixed in.

Figure 5.2. Team D1's Final written version of prompt 1



One aspect of the collaboration that was confirmed from the video was the relation between the grammatical LREs (LREg) that appeared in Extract 5.2 and the act of writing the final revised statement. At turn 27 when Monty other-corrected Eunice, she was writing. At turn 42, Mary was dictating a phrase for Eunice to write. In the next turn, Monty other-corrected Mary, adding in and stressing the word '*WHAT*' that Mary had omitted. The video confirmed that Eunice was writing at the time. In this way, the collaborative rewriting of VC task prompts promoted attention to grammatical form. In other words, when participants negotiated to a consensus about the linguistic content of their final version at the same time as they wrote it, LREg accompanied the rewriting process. This was observed in most of the other video recordings. The act of re-writing the prompt focused participants' attention to grammatical accuracy, a point I return to below (Section 5.4)

Task organisation

In Extract 5.2 task organisation occurred linguistically and through gesture. To introduce the new prompt, at turn 1, Eunice asked Mary, "*What do you think about number one?*" both nominating the topic and the interlocutor. Later, at turn 21, when the discussion moved from opinion to revision, there was no spoken signal. The only indication that they were now rewriting co-occurred with the first negotiation of member role. While Monty was speaking, Mary gestured to Eunice to write down what he was saying by holding her pen in the air and moving it as if she was writing (Fig. 5.3). As Mary was older than Eunice, important in Confucian societies, Eunice complied without saying anything. In the pause at turn 33, Mary pointed to Eunice to write on the answer sheet and not her prompt sheet (Fig. 5.4). They used words for the task, but for management, gestures served an important function for this team.

Figure 5.3. Mary using gestures to organise task roles- she tells Eunice to write



Figure 5.4. Mary using gestures to indicate where to write



5.2.2.3 Summary of the two teams' task performance

Extracts 5.1 and 5.2 demonstrate that both teams A5 and D1 have completed the VC task for prompt 1 of Lesson 5. They have each provided a final version of the prompt in writing (Figures 5.1 and 5.2). However, the process taken to complete the task differed. Team A5 focused on expressing their opinions and Team D1 focused their energy on revising the prompt to accurately reflect a collective opinion. Both teams were observed to be cooperative and friendly with each other, and neither team had a dominant member,

yet the former was more talkative. Neither team strayed from the task-as-workplan in a great manner, but neither did the task exactly as planned. Team A5 spent little to no time reaching a consensus, but spent a great deal of time discussing their opinions. In contrast, Team D1 was opposite, with most of their energy directed at creating a rewritten prompt.

5.3 Learning opportunities in the VC task

This section looks at the main ways the VC task provided learners the opportunity to develop their language and content knowledge. It looks first at collaboration, then at repetition and generation of language from the prompts, and finally at the mediating role of L1. This section draws on evidence from the task performances, whereas the next section draws on students' reflections on the task.

5.3.1 Learning through collaboration

As discussed in Chapter 2, Foster and Ohta (2005) defined four interactional processes through which learners collaborate to assist each other as they interact: co-construction, other correction, self-correction and continuers (Foster & Ohta, 2005, pp. 420–1). These four processes were richly evident in the data as we saw in Extracts 5.1 and 5.2. I will now discuss each in turn.

Co-construction occurs on two levels in Extract 5.2. The entire composition of the final version of the prompt, from turn 21–43 was an extended co-construction. Eunice, Monty and Mary did not once say the full statement individually, yet, at the end, it was written on their answer sheet. To do this, they combined their resources and created the final version by choosing a word (or words) at a time. There was limited discussion about these words, just suggestion and counter-suggestion until an agreeable expression was found. This series of smaller joint co-constructions created the final version. For example, from turns 23–25 Monty nominated '*has*', Mary supplied '*is bad for us*' and Monty supplied '*has bad things*'. This was not a case of Monty and Mary correcting each other. They were jointly supplying choices for the statement, creating a series of turns, each nominating potentially acceptable combinations, without evaluation or judgement. Similarly, they moved through '*bad things*' to '*bad points*' to '*bad aspects*' which became

their final choice (turns 33-38). This type of collaborative co-construction – a series of nominations for inclusion in the final re-written version of the prompt was used by teams throughout the study.

Self- and other-corrections about language are synonymous to LREs (Swain & Lapkin, 1995, 1998, 2001). In Extract 5.2, the only corrections were other-corrected LREs. It was noted above, that in Extract 5.2 the grammatical LREs (LREg) occurred while Eunice was writing the revised prompt (turns 27 and 43). In the first case, as Eunice wrote, she repeated what she was writing aloud and Monty corrected her. In the second case, Mary dictated an incorrect clause to Eunice, Monty overheard and corrected it. Their final version of the prompt (seen in Figure 5.2) was this correct version.

Other-corrections about vocabulary also appear in Extract 5.2. The vocabulary LRE (LREv) at turns 7-9 began with a gesture. Monty's hand movements combined with his inability to retrieve the word prompted Eunice's suggestion. Perhaps this is other-correction, but the process was 'softer', as Eunice's rising intonation implied her own uncertainty. Ohta's (2001) broader term 'assistance' seems more appropriate than 'correction' here as Eunice provided help for Monty so he could finish his idea. Just as it was noted that LREg's appeared while participants were writing, LREv's often appeared near the beginning of talk about a new prompt. In Extract 5.2, turns 6-10 illustrated this. This suggests that the participants experienced the VC task as a sequence with comprehension near the beginning of the talk and a move to accurate, written forms near the end.

Continuers were considered a form of assistance as they prompt a speaker to continue. The continuers used in turns 6, 10 and 17 all prompted another speaker to continue giving their opinion. This illustrated Team D1's cooperative nature, but none of the turns expressing opinion is either as long or as fully developed as those in Team A5 were.

Team A5's performance in Extract 5.1 shows a different focus of these interactional processes. There were no cases of co-construction of linguistic knowledge based on Foster and Ohta's (2005) definition, because Cathy, Lucky and Jessie were not focusing on language. There was only one self-correction for vocabulary (LREv). In turn 27, Lucky corrected '*way*' to '*side*'. The use of other-corrections highlighted this team's focus on content and meaning. Jessie used L1 to correct Cathy's interpretation of the prompt at turn 21, and Lucky explained the contents of the lecture for the other two after Jessie

asked (turn 31). Both were cases of a member needing to understand the meaning of the prompt, prior to continuing with the task as mentioned above. The members of Team A5 used the continuer '*really*' (turns 23 and 32) to stimulate further talk on their interlocutor's opinion.

Yet, despite the lack of co-construction and LREs, Extract 5.1 showed a successful VC task performance for one prompt. First, all three participants formed and expressed extended opinions and provided reasons for holding these opinions. From an SCT perspective, putting thought into words generates learning (Vygotsky, 1986). In addition, the supportive roles of waiting and listening to their interlocutors (Ohta, 2001) allowed the team to effectively revise the prompt. Jessie was able to make an acceptable version of the prompt on her first attempt because she had listened to the other two speakers and knew their opinions. She then restructured the prompt in a way that addressed Cathy's problem with '*too one-sided*' with Lucky's outright agreement with the original prompt. When seen this way, the VC task is as much a listening/input task as it is an output task.

5.3.2 Learning through mining the prompts

Another learning opportunity in the task was the use of language items, words and expressions, directly from the prompt in the task performance, what Boston (2008) called mining of 'task input'. This turned out to be extensive. The VC task required participants to read every prompt silently and indicate their level of agreement or disagreement with it. Sometimes they read the prompt again aloud to their partners (e.g. Extract 5.1, turn 57), and sometimes to themselves. Then participants had to discuss their opinions about the prompt and then change its wording to a mutually agreeable form (i.e. rewrite it) with the condition that they would later be asked to explain their changes.

The entire discourse of the VC task originates from understanding and altering the language of the prompt. Regardless of whether a team followed Team A5's style (Extract 5.1) and focussed on expressing their opinions in longer turns, or if they followed Team D1's style (Extract 5.2) and focussed on collaboratively generating a final version, any group of students must still confront and discuss the prompt. Mining of the prompt, then, was one way in which the VC task aided learning. Extract 5.3 exemplifies how a prompt's words and expressions are used and recycled during task performance. It shows Team B3, a team of three female students: Helen, Sara and Sue, discussing the prompt: "*It is not a good idea to eliminate gender biased language because language is too culturally*

important to change,” from Lesson 3. Words that are from this prompt are in **boldface** (even if the form is changed) and words from other prompts in the same task have been underlined.

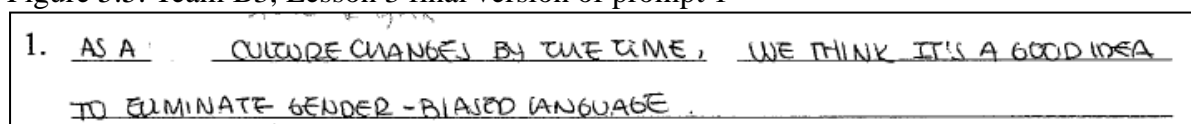
Extract 5.3. Team B3, Lesson 3, Helen, Sara and Sue, mining from the prompt

Turn	Speaker	Utterance
5	HELEN	uh what do you think about number one ?
6	SARA	(3.0) mm (4.0) what is eliminate (LREv)
7	HELEN	mm ?
8	SARA	eli: mi: nate
9	SUE	L1 <i>ɔ̃ʔoʔɛʔ eobs-aeda</i> (tr= remove, eliminate)
10	SARA	° L1 <i>ɔ̃ʔoʔɛʔ eobs-aeda</i> (tr= <i>remove, eliminate</i>) is not a good idea ° hmm strongly uh disagree ?
11	HELEN	yeah me too why ?
12	SUE	but here it says ° (<i>unint</i>) ° (most likely Sue is reading part of the prompt but the sound is obscured)
13	HELEN	I think that because we are EQUALS and about gender bias (/bi: as/) words something that uh before long time ago (2.0) ehh: (5.0) what do you think ?
14	SUE	hh the culture always changes by the . //time\\
15	HELEN	//years\\
16	SUE	yeah by the time . so: . now <u>women</u> - some womens (self-correction, incorrect LREg) are better than men so: ? . I strongly disagree
17	HELEN	but here says about language not about but who has more <u>power</u>
18	SUE	yeah mayb- because . if there are <u>more powerful women</u>
19	HELEN	mm-hmm
20	SUE	they have to call them
21	HELEN	ah:
22	SUE	but how are they going to call them ? if there's no language for <u>women</u>
23	HELEN	uh huh
24	SARA	ha
25	SUE	so they can- they uh I totally disagree
26	SARA	mm-hmm
27	HELEN	yeah yeah
28	SARA	mm:
29	HELEN	so how we change this sentence ? what do you think ? (to KSA)
30	SARA	uh . °what °
31	SUE	as a time passes culture changes it is a good idea to elimin- eliminate gender bias (Helen starts writing - video)

32	HELEN	suppose a- as (Helen requests Sue to repeat – gesture on video)
33	SUE	as a culture changes by the time
34	HELEN	changing
35	SUE	changes
36	HELEN	ah
37	SUE	by the time
38	HELEN	°mm-hmm ° yeah ?
39	SUE	I think it's a good //idea\\
40	HELEN	//we think\\ (LREg- other correction – while writing)
41	SUE	uh-huh we think it's a good idea ? (2.0) to: (4.0) eliminate (2.0) gender bias
42	HELEN	(6.0) h* mm language (2.0) okay then number two ?

Of the 12 content words (nouns, verbs, adjectives, adverbs) in the prompt, 10 different ones have been used by the participants. They have also used part of prompt 3 in this talk. The task has performed as expected. It has provided the participants the motivation to discuss both the topic and language, and has given them the opportunity to use and reflect on the language in the prompt. The more encounters a student has with a word or expression, the more likely they are to remember it (Joe, 1998; e.g. Webb, Newton, & Chang, 2013). In the case of Extract 5.3, if we include the original task prompt (written on their sheet) and the final written answer (Fig. 5.5), the VC task shown in Extract 5.3 provided five exposures each to the formulas: ‘*gender bias*’ and ‘*a good idea*’.

Figure 5.5. Team B3, Lesson 3 final version of prompt 1



A look back at Extracts 5.1 and 5.2, above, also reveals a great deal of mining and recycling of the language in the prompts, because task performance depended on it. Mining was a task-essential process (cf. Loschky & Bley-Vroman, 1993). The task could not be done without it. The mining and repetitive use of the same items multiple times ‘disguised’ in an opinion gap, was one way, in conjunction with the collaborative interactional features of the task that learners acquired and consolidated language during the VC task.

Although the language of the prompts in this study was not pre-specified along lexicogrammatical lines, the task provided a medium for participants to use and reflect on the

words and expressions in the prompts. This language was ‘task-essential’ (Loschky & Bley-Vroman, 1993) and therefore could not be ignored. I did not use linguistic specifications for the prompts in this study because I wanted the VC task to be based on the content information objectives of the course. The students task performance of the VC show that there is room for including lexico-grammatical objectives within the content objectives of the VC task. It will be argued later that in other learning situations, prompts could be seeded with vocabulary or grammar items that reflect local institutional demands (Conclusions, Chapter 9).

5.3.3 Learning of the course material

Another learning opportunity afforded students by the task was the consolidation of course content. This study took place in existing classrooms with predefined learning objectives. The study was part of a regularly administered sustained content language teaching (SCLT) programme. In this context, implementing tasks that had little or no relevance to course content could be problematic for students and administrators. The values clarification (VC) task was flexible enough to be adapted to any topic about which an opinion can be argued. Extracts 5.1 and 5.2 illustrate the relationship of the task to the course content. The prompt referred directly to the text book material, but Jessie, in turn 19 stated that the material was new. Later, she directly asked about it, and it turned out that Cathy did not know for certain either, leaving it up to Lucky to explain (turns 29-33). In this way, they reviewed course content through the task, and in so doing, highlighted the potential of the task as a vehicle for content learning (cf. Brinton, 2003).

5.3.4 Summary of learning in the VC task

This section illustrated the processes by which students’ choice of approach to the values clarification (VC) task provided different opportunities for language and content knowledge development. Interactive collaboration provided the basis for co-constructed utterances, self-corrections, and other corrections (the last two items equating with LREs). It was shown that participants used continuers to keep their partners talking about their opinions, thereby keeping them using language and engaged with language. The language of the prompts was shown to be mined heavily as a result of the task design. This involved multiple repetitions of items across written and spoken modes. The final change of mode from speech to writing was shown to generate grammatical LREs, and bears further examination. In contrast, the initial change of mode – from reading the prompt, to

speech, done at the beginning of discussion on a prompt, appeared to generate more lexical LREs. Finally, due to the topical relationship of the task prompts with the course content, the VC task provided an opportunity to review and consolidate content knowledge.

These findings were based on data from class audio and video recordings and my observation notes. The next section offers the participants' perspective on learning in the VC task, and confirms much of what was claimed here.

5.4 Student voices about learning from the VC task

The students' perspectives on their task performances and learning experience in the programme were vital to the internal validity and overall reliability of the findings. Their words explain their performance from their perspective. This emic data grounds the study in the students' experience, not in my, teacher-researcher's, mediated view of it. My analysis is thereby framed in theirs, lending greater ecological validity to the findings. Their ideas also provided me with a foil to challenge my interpretations based on my assumptions about teaching, learning and classroom roles, and provided me with a blueprint to improving classroom learning opportunities in ways I had not thought of (See Chapter 6).

To investigate their thoughts, I gathered data from two sources. I interviewed ten students at the end of the study (see Ch. 4, and Appendix 5) and, I also had students perform a student-student survey as the central element of the MA activity (see Ch. 4 and 8). Because the MA intervention was a key task implementation condition, it was studied in its own right, for its effects on subsequent task performance. However, due to the nature of the MA survey which asked about how students did the task, what they did when encountering difficult input and so on, the students' responses also served as a means to investigate how students approached the VC task and how they felt about performing it multiple times.

In addition, the MA activity controlled for any response bias in the teacher-student interviews. I was concerned that the power relationships involved in an instructor-led

interview, albeit one conducted after the semester had ended, may still make students hesitate to express negative opinions. I felt that they might feel freer to express any opinion when talking to a team of peers. Investigating the MA activity also allowed me to get the opinions of all students and not only the ten who volunteered for post-study interviews, thereby eliminating another source of sampling bias. I will show in Chapter 8 that the students reacted differently to what their peers said than to what the teacher said.

From both sources, I found that in general, the students had realised that the VC task was implemented to help them develop their language skills and to consolidate the information in the course textbook. I turn first to their ideas on language development, and then move to their ideas about content consolidation second.

5.4.1 The VC task and language development

The participants reported and discussed a variety of ways that the VC task aided their language and language skills development.

5.4.1.1 The VC task and vocabulary development

All interview participants mentioned that the VC task was good for their vocabulary development. In Extract 5.4, Jessie recognised that there were two parts of a VC task: to express an opinion and to re-write the prompt. She felt the two stages were helpful and that the VC task was better than a discussion task in developing her vocabulary.

Extract 5.4. Jessie saying why the VC task is helpful for language development

Jessie

Yes I think that's [the VC task's] helpful. And also if we just choose my opinion and just talk our- MY opinion then it will less helpful, but we have to change that sentence and we have to make a same opinion, so...

Researcher

How does that help?

Jessie

Uh. If I say only my opinion then then I have a limit to using a word or expression but if I heard another person's opinion and we have to add and make a new sentence then I have to use new expression or new word, so it's very helpful

Her contrasting of the VC task with a discussion task is illuminating because she considered listening to others and reaching a consensus to be important in developing her vocabulary. In her opinion, because of the rewriting, as she put it, “*change that sentence and ...make a same opinion,*” she could expand the limits of her vocabulary use. Making the new sentence meant that Jessie had to “*use [the] new expression*” which she stressed was helpful for learning. In this comment she did two things: she stressed the role of peer-assisted learning, and the role of rewording the statement. The latter drives the communicative need for new words, and the former is the source of the new words.

During group discussion talk in the metacognitive awareness (MA) intervention, Jessie repeated this idea to her group members.

Extract 5.5. Jessie claims that adjusting the sentence means using new words

if we just talk about my opinion then always say my opinion.. I just use my words maybe one or ten word ... if we make a new sentence we use new word and ...make a new sentence for us make us use more new words or grammar.

By using the expression “[if] *I just use my words*” Jessie indicated she needed others to supply the new words, which they collectively use when rewriting the statement. She also realised that writing the new sentence is a form of learning new words because it “*make us use more new words or grammar*”. This highlights her view of the importance of revising the sentence to language development. Extract 5.5 added grammar development to vocabulary mentioned in Extract 5.4. This is echoed by others, below.

5.4.1.2 The VC task and grammar development

Many other participants reported that grammar developed because of rewriting the prompt. Lucky, for example, claimed that rewriting the prompt helped her improve her grammar, through noticing her mistakes. When she was asked about the purpose of the rewritten, final version of the prompt she answered the following:

Extract 5.6. Lucky commenting on writing in the VC task

I guess it's part of grammatic problem- problems I guess...because when we are speaking we don't know what our speaking is- how I mean HOW our speaking is wrong like grammatical order or other problems but when we WRITE that sentence

in the paper then we notice what is wrong or what is right

This extract shows that Lucky considered the act of writing the final version as a kind of feedback, a way to notice flaws in her production that could not be noticed through speech. Her comment also indicated that she felt unable to notice mistakes of form during speech. This does not mean that the participants could notice all the grammatical errors in their writing, but the move to a written form, according to Lucky, aided attention to form. This corroborates the anecdotal finding above, where it was noted that many LREg occurred while one of the team members was writing their final version of the prompt.

5.4.1.3 The VC task and comprehension skills development

One of the stages of the VC task described above was comprehension of the prompt. Some participants reported that the VC task aided their comprehension skills. In Extract 5.7, Steve commented on the need to understand the prompt as a prerequisite to agreeing or disagreeing with it.

Extract 5.7. Steve talking about the role of comprehension in the VC task

We did the task that change the sentence so we can make our opinion. Just to change the sentence we have to know the sentence

Here, Steve referred to understanding the meaning of the prompt. The first learning opportunity afforded by the VC task is comprehension of the prompt. It was observed above that LREs sometimes clustered near the beginning of a discussion about a new prompt, because participants did not fully understand the language of the prompt. We shall see in Chapter 7, that even when LREs appear to be resolved, there may still be comprehension problems requiring further language work. Steve's point, here, suggests there may be learning potential in seeding the prompt with targeted lexico-grammatical items instructors or institutions expect their students to learn.

5.4.1.4 The VC task and speaking skill development

The participants were first year students at a Korean university and as such had received very little opportunity to say anything in English, let alone give their opinion, before

taking this course (Chapter 1). Sunny, in the extract below, stated that the VC task was helpful in that it aided her speaking development.

Extract 5.8. Sunny's statement on the relation of the VC task to school learning

I also absolutely [think] it is helpful, because we we just studied only English reading or grammar I can speak English ...not fluently but it is uh step by step ...

She emphasised that the VC task helped her fluency develop “*step by step*”, implying the need for repeating the task.

5.4.1.5 The VC task and learning through peer interaction

Jessie also mentioned the role of peers in learning (Extracts 5.4 and 5.5). Winnie echoed this. Here, she identified herself as one of the weaker participants in the class. She felt that when she interacted with her group members, she could learn from them just through hearing their expressions.

Extract 5.9. Winnie on learning from more advanced peers

because uh ... almost [all] people in this class can speak better than me so when I talk with them they use English expressions that I don't know so I can learn from them

Winnie referred to the process of doing the task and learning during the talk from more advanced speakers. Her view was the task performance afforded her learning opportunities by providing new expressions from the input of her group members.

5.4.2 The VC task and content-knowledge development

This course was a sustained content language teaching course (SCLT) (Brinton, 2003). Its primary objective was to equip students in the course with the skills needed to participate successfully in English-medium classes offered in their respective majors at that university. To achieve this objective, both the English proficiency of the participants and their mastery of content knowledge were assessed for grades.

In the metacognitive awareness (MA) intervention (see Chapter 8), participants frequently

referred to this objective when discussing the purpose of the VC task. They considered it a means of consolidating content knowledge and checking for understanding of the previous lessons, as Jessie claimed in Extract 5.10.

Extract 5.10. Jessie's view of the VC task as content knowledge consolidation

...it is his [the instructor's] objective. If we talk about from book's topic we can arrange our opinion and compare with this book ... we can remember better ... every time we need our opinion and can remember better

Here Jessie claimed that making and using their opinion develops their ability to remember the content better. She viewed the VC task as part of the process of learning content. This view was echoed by Steve, in Extract 5.11.

Extract 5.11. Steve's view that VC talk as a comprehension check of content

He [the instructor] wants we know the topic ...making our opinions means we understand the book's topics so it is his intention.

Jessie and Steve both referred to the content of the course, not English, as being the primary motivation for doing VC tasks. In Jessie's view, the VC task was used to help them remember the content from their textbook, a kind of knowledge consolidation and review. Steve's opinion mirrors hers. For him, the VC task was the teacher's way of making sure they knew the material by making them give their opinion about it. Considering that the prompt discussed in Extracts 5.1 and 5.2, above, directly required participants to remember the previous lesson's contents, the point of view expressed by Steve and Jessie was realistic and echoed my reasoning for using VC tasks. The VC task was chosen because it integrated easily with the topical content of the course. It did this so well that the author of the course textbook (Sanabria, 2004) later used a VC task in a different title in the same series of books (Sanabria & Sanabria, 2008, p. 119).

5.5 Conclusions about the VC task

The VC task is based on a gap in opinion. To generate the opinion gap, participants had to understand the prompts, both on a linguistic level and on a course-content level. This effort to understand often led to LREv's. The expression of opinion and construction of a rewritten opinion typically took one of two paths. An opinion-driven path, exemplified in Extract 5.1 or a rewriting-driven path, exemplified in Extract 5.2. LREg's tended to appear while participants were writing the final version of their prompt, due to increased ability to notice what was being produced according to one participant. Another participant praised the VC task with respect to discussion tasks because it forced her to use the words of other people and adapt her opinion with them. The requirement for rewritten prompts therefore pushed vocabulary development as well.

This shows how the move across modes, from speech to writing, pushes lexico-grammatical awareness. This is not possible in oral tasks and integrated speaking-writing tasks are not common. One well-studied exception, however, is the dictogloss task (Wajnryb, 1988). Simply put, in the dictogloss learners listen to a dictation and take notes. They then collaborate to reconstruct the original. However, its 'gap' is essentially a 'proficiency' gap; differences in what was heard during the dictation phase of the task drive the talk in the reconstruction phase. When contrasted with the opinion gap that drives the interaction in VC task, a picture of the different mental processes involved emerges. The words and ideas of the materials developer are central to the dictogloss whereas the words and ideas of the participants are central to the VC task. Therefore, dictogloss involves aural perception and recognition, the VC involves recall and synthesis of topical information.

Both tasks can be used as focused tasks (Ellis, 2003) – they allow for seeding the input with targeted lexico-grammatical structures or to support a content-based course, but only one, the dictogloss, has been researched to any great extent. It is argued here that the values clarification task deserves closer examination to assess its potential as an aid for language learning.

5.6 Chapter Summary

This chapter reported an ecologically valid examination of values clarification (VC) task performance, and discussed how VC task performance involved different phases in which different performance processes were used. The VC task involved four phases. Three phases required the use of language resources, i.e. comprehension of the prompt, expression of opinion, and rewriting a consensus version of the prompt that all team members 'strongly agreed' with. The fourth, organisation of their performance, was shown to involve both language and gesture. The teams presented in this chapter engaged the phases of the task in different ways, yet reached successful conclusions. Although each team emphasised different phases, they still followed the task-as-workplan to a large degree. The workplan itself required a great deal of interaction and a focus on the language of the prompts. What differed between the teams was the interactional focus: opinions or rewriting the prompt, i.e. opinion gap or task completion.

Interaction in the VC task was highly collaborative. In one of the teams presented in the chapter, a great deal of co-construction took place, along with LREs (self and other-corrections), and the use of continuers to prompt the interlocutor to continue speaking. In the other team, there was more interactive talk about task organisation and about the meaning of the prompt. When dealing with opinions, that team took a wait and listen approach, allowing each member to have longer turns with more complete thoughts.

Due to the nature of the VC task, all teams always needed to mine the prompt they were talking about for language (Boston, 2008). The participants not only had to discuss the topic, they had to revise the prompt's language. This latter component led to a great deal of repetition of words and expressions from the prompt in the participants' VC talk. Although not a feature of this study, seeding these prompts with targeted vocabulary or grammar structures appears to be one way of integrating a communicative task into an existing language classroom.

In addition to seeding the prompts for language, they can be used for consolidating content knowledge. This was one of the purposes for using the VC task in this classroom's context. In interviews and metacognitive awareness (MA) survey interaction, participants identified content consolidation as one of the purposes I had in mind when

assigning this task. The participating students also stated they learned vocabulary and grammar because of the VC task's two-part structure: discuss – then – revise. In short, they experienced the task in the way I expected, lending credence to my claim for the ecological validity of the study (Bronfenbrenner, 1979; Van Lier, 2004).

This chapter has provided a look at how the VC task was carried out by the participants: their task-in-process. As such, it serves to ground the remaining three chapters of the results and discussion. The next three chapters report the results of the quasi-experimental interventions done in this second phase of the research. Chapter 6 presents the impact of an input processing activity on VC task performance. Chapter 7 presents the impact of task repetition on VC task performance; and Chapter 8 presents the impact of a pre-task metacognitive awareness (MA) raising task on VC task performance.

CHAPTER SIX: THE EFFECT OF AN INPUT PROCESSING ACTIVITY ON TASK PERFORMANCE

6.1 Introduction

This is the second of four results and discussion chapters. The previous chapter provided an overview of the values clarification (VC) task. This chapter examines the effects of the placement of an input-processing activity: pre or post-task. First, it briefly restates the research design and research questions. Then it reports the quantitative results, followed by qualitative data that helps explain the results. Next, the findings are discussed in relation to other studies and theoretical perspectives. The chapter concludes with a summary of the main findings.

6.1.1 Review of the research design

The first four lessons of the study were used for this investigation. Lessons 1 and 2 used the same VC task (VC₁): an original version followed by a repeated (and modified) version. Lessons 3 and 4 used a different VC task (VC₂); also with an original version and a repeated/modified version⁴.

Table 6.1 shows the arrangement of lesson, topic, task and input placement condition. Group 1 consisted of participants from two classes: A and D; and Group 2 contained participants from the other two classes, B and C. Two classes for each condition were needed to provide enough overall participants. Topics were determined by the chapters of the course textbook being used at the time of the VC tasks (Sanabria, 2004).

The input-processing activity consisted of: (1) listening to native speakers doing the same task; (2) reading a transcript of their interaction; (3) answering comprehension questions about it; (4) doing consciousness raising activities about some of the items from the script; and finally (5) doing a brief set of repetition drills about a few of these targeted items.

⁴ Two versions of each task were used in order to examine the effects of task and procedural repetition (Kim, 2013) discussed in the next chapter. See Chapter 4 for full methodology.

Table 6.1. Arrangement of lesson, topic, task, and input-placement sequence

Lesson	Topic	Task	Input-placement condition sequence	
			Group 1 (N=22)	Group 2 (N=24)
1	Single-sex education for girls	VC ₁ original	task-then-input	input-then-task
2		VC ₁ repeated	no additional input-related activity	
3	Gender and language	VC ₂ original	input-then-task	task-then-input
4		VC ₂ repeated	no additional input-related activity	

6.1.2 Research questions

The research questions addressed in this chapter are as follows:

1. Does placement of an input processing activity (IPA) before or after the values clarification (VC) task affect task performance?
 - a. Does the number of LREs change with IPA placement?
 - b. Does the number of turns of L1 use change with IPA placement?
 - c. Does the use of target items from the IPA activity change with IPA placement?
 - d. Does the target-like use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993) change with IPA placement?
 - e. Are the changes found in both original and repeated (delayed) task performances?

6.2 Quantitative data and analysis

To measure the effects of the input placement activity (IPA) on task performance, the dependent variables of LREs (number and type), turns of L1 use, the number of items from the IPA's language awareness activity used in the task performance, and the number of target-like uses of the phrasal verb '*agree with*' were used as dependent variables (See Chapter 4).

6.2.1 The effect of the input-processing activity on LREs

The first dependent variable investigate was the number and type of LREs. Language

related episodes (LREs) are taken as evidence of learning in interaction (Swain & Lapkin, 1995, 1998). An increase in the number of LREs indicates more attention being paid to aspects of language the learners had not yet mastered. It was expected that students in the task-then-input condition would face more language problems and hence would produce more LREs as they had not had any input to learn from to help their output.

Because the original VC tasks had five prompts and took more time than the repeated versions, all data had to be standardised to a value per 10 minutes of task performance. Table 6.2 contains the standardised data for LREs⁵. The shaded areas in the table indicate the participants from Group 1 and the non-shaded data indicates the participants from Group 2. An examination of Table 6.2 reveals that there is little difference in grammatical LREs (LREg), vocabulary LREs (LREv), pronunciation LREs (LREp) or Total LREs across the two conditions. There were a total of 137.4 LREs for the task-then-input condition and 140.4 in the input-then-task. Similarly, LREg's and LREv's were marginally higher in the input-then-task condition.

In order to confirm this observation, independent samples, Mann-Whitney U tests were conducted on total LREs, LREg, LREv, and LREp due to the small number of teams in each condition (N=7 per condition). The null hypothesis claimed no difference between input-then-task and task-then-input conditions. The results of the Mann-Whitney U tests for total LREs are included in Table 6.3. For all four variables, no statistically significant differences were found and the null hypothesis was retained. The finding, therefore, was not as predicted; there was no difference in total LREs, grammatical LREs, vocabulary LREs, or pronunciation LREs across input conditions. Thus, the pre-task input-processing activity had little or no effect on the number or type of LREs produced in the VC task talk. This finding is presented visually in Figure 6.1.

⁵ The raw data, uncorrected for time, appear in Appendix 9.

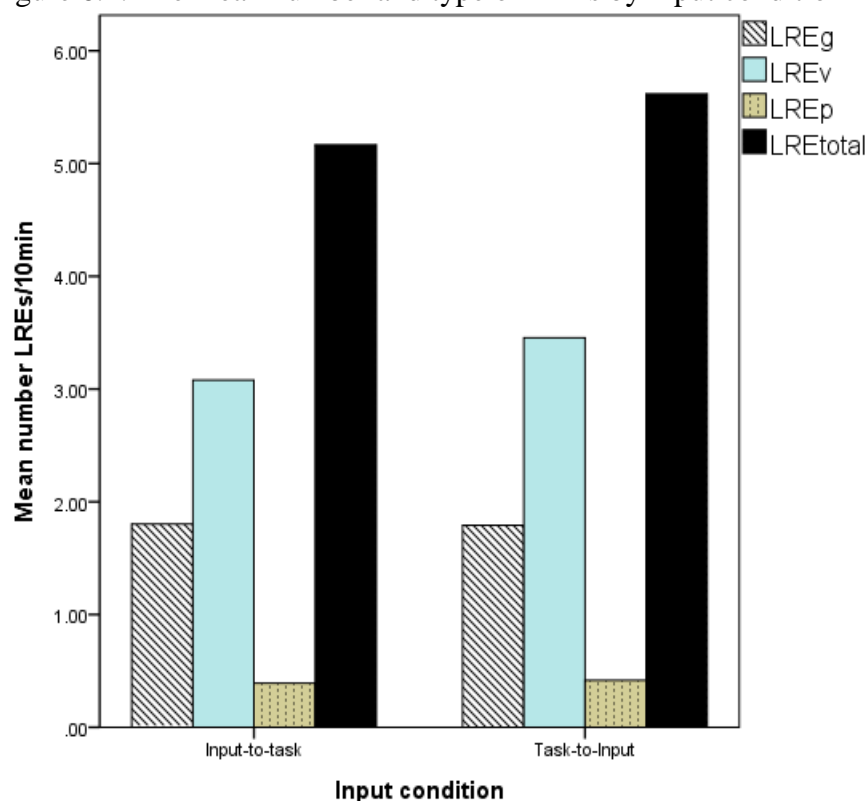
Table 6.2. The number of LREs per 10 minutes of task performance across input-placement condition

Lesson	Topic	Task	Input-placement condition sequence								GRAND TOTAL
			Task-then-input				Input-then-task				
			LREg	LREv	LREp	TOTAL	LREg	LREv	LREp	TOTAL	
			Group 1				Group 2				
1	Single-sex education for girls	VC ₁ original	10.4	27.3	3.4	41.1	14.3	27.1	3.6	45.0	86.1
2		VC ₁ repeated	13.7	21.3	0.9	36.0	9.2	20.5	0	29.7	65.7
Subtotal 1: LREs used in VC ₁			24.1	48.6	4.3	77.1	23.5	47.6	3.6	74.7	151.8
			Group 2				Group 1				
3	Gender and language	VC ₂ original	10.3	19.7	4.9	34.8	12.1	17.2	3.6	32.9	67.7
4		VC ₂ repeated	6.3	18.2	1.0	25.5	6.6	23.4	2.8	32.8	58.3
Subtotal 2: LREs used in VC ₂			16.6	37.9	5.9	60.3	18.7	40.6	6.4	65.7	126.0
GRAND TOTAL LREs			40.7	86.5	10.2	137.4	42.2	88.2	10	140.4	277.8

Table 6.3. Results of Mann-Whitney U tests for total LREs per 10min task talk across input condition

Task	Input condition	Mean total LREs / 10min	Mann-Whitney U statistic	Asymp. Sig
VC ₁ original	task-then-input	5.8	25.00	0.948
	input-then-task	6.3		
VC ₁ repeated	task-then-input	5.3	28.50	0.607
	input-then-task	4.2		
VC ₂ original	task-then-input	5.1	22.00	0.747
	input-then-task	4.6		
VC ₂ repeated	task-then-input	3.6	32.00	0.332
	input-then-task	4.7		

Figure 6.1. The mean number and type of LREs by input condition



As the statistical tests confirm (Table 6.3), there was no difference between Group 1 and Group 2 for input condition. The rows of subtotals for LREs in VC₁ and VC₂ show a small tendency for Group 1 to have more LREs than Group 2, but this was not statistically significant.

The effects of task repetition will be discussed in the next chapter, however, there was an overall decrease in LREs over time across the four lessons (i.e. VC₁ original, VC₁ repeated, VC₂ original, and VC₂ repeated) in both input-then-task and task-then-input conditions.

6.2.2 The effect of the input-processing activity on L1 use

For the original version of the task, it was predicted that providing a pre-task input-processing activity would reduce the amount of L1 used by the groups because the model provided lexis, grammar and ideas for the participants to refer to (See Ch.4). In a repeated version of the task, it was predicted that L1 use would decrease vis-à-vis the original in both conditions, but there would be a greater reduction in the task-then-input condition due to more L1 use predicted in the original task.

An L1 use turn was defined as any turn using the participants' L1 (Korean) except for: (1) proper nouns (e.g. place names), (2) LREs mediated with L1 as these had been categorised

elsewhere, and (3) in the case of VC₂, Korean terms used as examples of gender-biased language and integrated into otherwise English utterances. On-topic L1 was defined as uses of Korean that related to the prompt under discussion, provided (or asked for) additional background information, re-phrased (or requested a re-phrasing) English utterances, or that managed roles and turns in the interaction. Off-topic L1 was any utterance in Korean that related to something other than the prompts, input or task performance. On and off-topic L1 use were distinguished because L1 on-topic use was considered a potentially useful resource for language learning, whereas off-topic L1 was considered unhelpful.

Table 6.4 gives the number of turns in which L1 was used per 10 minutes of task talk. The Grand Totals in the bottom row reveal the following: (1) slightly more total L1 use in the task-then-input condition than in the input-then-task condition (241 turns vs. 235 turns); (2) Less on-topic L1 in the task-then-input condition (149 vs. 162); and (3) more off topic L1 in the task-then-input condition (92 vs.73). Because the L1 data was heavily skewed (skewness of over 3.0 in some cases) and kurtotic (kurtosis of over 11.0 in some cases), parametric statistics could not be used. Mann-Whitney U tests were performed and the results appear in Table 6.5. No statistically significant differences for total number of L1 turns, the number of on-topic L1 turns or the number of off- topic L1 turns were found. The graphic representation of this data shown in Figure 6.2 underscores this finding.

This was not unexpected, as it had been predicted that when participants received pre-task input they would not use as much L1 because the input provided a model for the interaction and samples of language. Yet the overall difference in total L1 across task-then-input and input-then-task conditions was not statistically significant, only about 5 turns per 10 min of talk. For comparison, the overall mean number of turns across all groups and all tasks was 153 per 10 min.

The overall decrease in L1 is somewhat misleading, however, because the prediction could only be made with respect to on-topic L1. However, on-topic L1 use was lower in the task-then-input condition, counter to predictions. Again, this result was not statistically significant (see Table 6.5), but it does raise questions about the way the input processing activity influenced L1 use. We have already seen in the previous section that the activity had no effect on LREs. With respect to L1 use, the claim that the input-processing activity had no effect cannot be rejected.

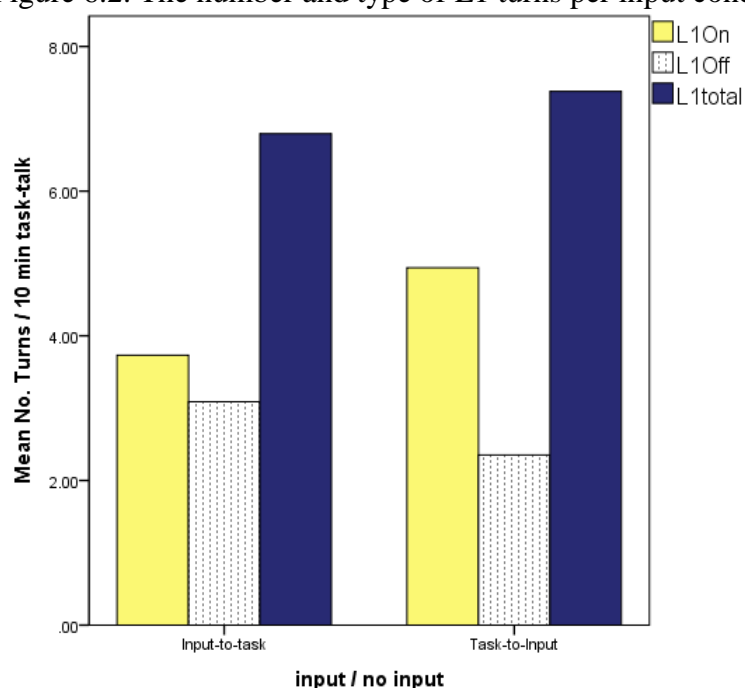
Table 6.4. The total number of turns using L1 per 10 minutes of task performance talk

Lesson	Topic	Task	Input-placement condition sequence						TOTAL
			Task-then-input			Input-then-task			
			L1-off topic	L1-on topic	Total L1	L1-off topic	L1-on topic	Total L1	
			Group 1			Group 2			
1	Single-sex education for girls	VC ₁ original	4.9	67.8	72.7	5.1	21.6	26.7	99.4
2		VC ₁ repeated	37.4	40.9	78.3	13.0	34.9	47.9	126.2
Subtotal 1: L1 use for VC ₁			42.3	108.7	151.0	18.1	56.5	74.6	225.6
			Group 2			Group 1			
3	Gender and language	VC ₂ original	12.5	27.4	39.9	31.4	46.5	77.9	117.8
4		VC ₂ repeated	37.0	13.3	50.3	23.8	59.4	83.2	133.5
Subtotal 2: L1 use for VC ₂			49.5	40.7	90.2	55.2	105.9	161.1	251.3
GRAND TOTAL			91.8	149.4	241.2	73.3	162.4	235.7	476.9

Table 6.5. Results of Mann-Whitney tests for L1 use across Input Condition

Variable	Mann-Whitney U statistic	Asymp. Sig
L1 total	370.0	0.716
L1-on	471.0	0.156
L1-off	368.5	0.699

Figure 6.2. The number and type of L1 turns per input condition



A potential explanation for this non-result might be the characteristics of the groups or the teams. Further inspection of Table 6.4, above, indicated that there may be a difference in the tendency to use L1 between Groups 1 and 2. The Total L1 column reveals that the least L1 used by Group 1 was 73 turns, in Lesson 1. The most L1 used by Group 2 was 50 turns, in Lesson 4. Likewise, each subtotal row in Table 6.4 ('Subtotal 1: L1 use for VC₁' and 'Subtotal 2: L1 use for VC₂') indicates more L1 use by Group 1 than Group 2 in all cells. However, Table 6.6 shows no statistical differences in L1 total, L1-on or L1-off due to differences in membership of each group.

Table 6.6. Results of Mann-Whitney tests for L1 across groups (Group 1 vs. Group 2)

Variable	Mann-Whitney U statistic	Asymp. Sig
L1 total	323.5	0.259
L1-on	303.0	0.141
L1-off	322.5	0.211

An examination of L1 use across input condition for original and repeated tasks was carried out to determine if the combining of original and repeated tasks hid any effect of the input-processing activity on either the original task or the repeated task. For the original tasks,

teams going from input-then-task were expected to use less L1. For the repeated tasks, task-then-input was expected to show a greater reduction in L1 use, but both conditions were predicted to show overall reductions in L1 use. Table 6.4, above, appears to contradict this prediction. In all cases, the total L1 per 10 min talk increased over the original tasks. On and off-topic L1 use fluctuated in different directions. In three cases, off-topic L1 rose from original to repeated, and in one case it decreased. On-topic L1 rose in two cases and decreased in two cases.

Table 6.7 shows the results of the Mann-Whitney U tests run on the total L1 use data. Total L1 was compared across input condition for each VC task. For all variables, including on-topic L1 and off-topic L1 – not shown in the table, the tests were not statistically significant.

Table 6.7. Results of Mann-Whitney U tests for Total L1 turns per 10 min

Task	Input condition	Mean L1 turns / 10min	Mann-Whitney U statistic	Asymp. Sig
VC ₁ original	task-then-input	10.4	38.5	0.070
	input-then-task	3.7		
VC ₁ repeated	task-then-input	11.2	26.5	0.805
	input-then-task	6.9		
VC ₂ original	task-then-input	5.7	35.0	0.179
	input-then-task	11.1		
VC ₂ repeated	task-then-input	7.2	19.0	0.439
	input-then-task	11.9		

Finally, I investigated the extent to which L1 use varied among the teams. Table 6.8 shows the L1 use by each team in each condition. The right hand column, ‘Mean L1 turns per team’, demonstrates the variation across teams. Team A3 used the most L1, about 16 turns per 10 min. Team B2 used almost no L1, less than one turn per 10 minutes. Four teams used more than 10 turns per 10 minutes: Teams A3, C2, C3 and D1. Four teams used two turns or fewer per 10 min: Teams A5, B2, C1 and C4. Group 1 was made of Teams labelled A and D, and Group 2 consisted of Teams labelled B and C. Two teams from each group were in the ‘heavy’ L1 user category and three from Group 2 and one from Group 1 was in the ‘light’ L1 user category, indicating a relatively even distribution of high and low L1-using teams across groups.

Table 6.8 also shows that the difference between task-then-input and input-then-task

conditions for Total L1 is explained by the idiosyncratic nature of L1 use at the level of team. Some teams, such as, A2 and C2 used far more L1 in the task-then-input condition. Others, such as A4 and D2 used more in the input-then-task condition. And still others, such as A1 and C1, were relatively even across conditions.

Table 6.8. L1 use for each team (expressed in number of turns per 10 minutes of talk)

Team	On-topic L1		Off-topic L1		Total L1			Mean L1 turns per team
	Task-then-input	Input-then-task	Task-then-input	Input-then-task	Task-then-input	Input-then-task	Difference	
A1	2.5	2.5	0.5	2	3.5	4.5	-1.0	4
A2	6	5	19	3	25	8	17	16.5
A3	23.5	10	0.5	15.5	24.5	25.5	-1.0	25
A4	1	3	0	5.5	1	8	-7.0	4.5
A5	2	1	0.5	0.5	2.5	1.5	1.0	2
B1	1	2	8	4.5	9	6.5	2.5	7.75
B2	0	0.5	0	0.5	0	1	-1.0	0.5
B3	14.5	8.5	0.5	0.5	15	9	6.0	12
C1	2	1	0	0	2	1	1.0	1.5
C2	4	2.5	16	0.5	20.5	2.5	18.0	11.5
C3	14	12.5	0	2	14	14.5	-0.5	14.25
C4	1	0.5	0	1.5	1	2	-1.0	1.5
D1	14.5	8.5	0	0.5	14.5	9	5.5	11.75
D2	4	6.5	0.5	1	4.5	7.5	-3.0	6
Total	90	64	45.5	37.5	137	100.5	36.5	118.75
Mean per condition	6.43	4.57	3.25	2.68	9.79	7.18	2.61	8.48

This suggests the individual dynamics in a team influenced L1 use and not the input processing condition. This agrees with the findings of Storch (2002) who identified different interaction patterns in pairs. Here one key difference among teams (not pairs) was the amount of L1 they used. This will be addressed later in this chapter through interviews and other qualitative data.

6.2.3 The effect of the input processing activity on use of targeted items

The Willis' framework for TBLT included a post-task language focus (in this study called language awareness, LA) designed to address items the learners might not know. The language focus had two main parts. First was a series of consciousness raising (CR) or similar activities that focused learners' attention to language. They proposed basing this language focus on a script of native speakers doing the same task. The second part of the language

focus was a brief controlled practice, i.e. repetition drills, of some of the targeted items. They predicted that placing this language focus *after* the main task would be most beneficial for learners because it allowed them to notice the gap in their productive ability and therefore be primed to learn from the post-task. They also argued that placing the language work *before* the main task would lead to learners over-emphasising these items at the expense of correctly articulating what they meant to say (D. Willis & Willis, 1996; J. Willis & Willis, 2007; J. Willis, 1996a, 1996b).

This study set out to determine if this was the case. Willis' position predicted, that more items from the language awareness (LA) activity would be used and that when items were used, they would be used more correctly in the task-then-input condition than in the input-then-task condition. Obviously, no items could be used in the task-then-input condition, because students could not use what they had not yet seen. (Any items that were used would reflect previous knowledge.) However, the repeated version of the task would show differences if input-placement had an effect.

Surprisingly, there was only one example of participants directly and accurately using the targeted items from the language awareness (LA) activity in the entire data set. This unpredicted result happened despite the participants doing the LA activity with their group members and having the answers given and explained by the instructor. It was impossible for the items not to be noticed, yet they were not used.

The qualitative data in the next section will address the reasons for the lack of use of these items. However, I include here the one LA item that was used by a student.

The task performance of team A3 in Lesson 6 is shown in Extract 6.1. Nancy said “*range of her topic*”. The expression ‘[*past / beyond*] the *range of [sth.]*’ was the answer to Question 4 from Lesson 1, done six weeks prior to this task performance. Subsequent re-analysis of the data showed she had also used the term ‘*range of*’ in Lesson 4. It was not precisely used as she omitted the article each time, but it was verified that this was a new item from the LA activity and not known previously because recordings revealed that her team had not answered it correctly in Lesson 1.

Extract 6.1. Nancy using the item ‘*the range of*’

Turn	Speaker	Utterance
59	NANCY	what she wants what she said was very accurate . what she said was accurate uh ah <u>range</u> of her topic- topic was bad side so in that <u>range</u> it was accurate (5.0) uh her topic was:
60	SUNNY	uh
61	NANCY	bad side of mass media
62	WINNIE	yes yes

This item showed up twice, spoken by the same speaker, in lessons long after the item was included in the LA activity. It is probably an instance of delayed uptake, but we cannot rule out the possibility that she learnt it elsewhere in the meantime.

In the previous chapter, it was noted that there was a great deal of mining of language items in the VC prompts. In contrast, there was almost no use of LA items in the VC task performances. This finding will be addressed in Section 6.3.

6.2.4 The effect of the input processing activity on the correct use of *agree with*

The phrasal verb *agree with* followed by a noun phrase (NP) was chosen for analysis because VC task data from the exploratory (phase 1) study revealed that in situations when ‘agree’ was followed by a NP, it was almost always accompanied with the preposition ‘to’ and not ‘with’ (i.e. ‘Agree+to+NP,’). Most of the time, the participants said something similar to, “*I agree to your opinion,*” rather than “*I agree with your opinion.*” In the VC task, ‘agree with’ is a “task natural” form (Loschky & Bley-Vroman, 1993) as when doing a VC task, it is possible for participants to say “I agree” or “I disagree” without a following NP. Another, incorrect form had also been observed in the exploratory study: ‘BE+agree’ (e.g. “I’m agree”, “Why are you disagree?”). This was eliminated from further examination in the intervention study (phase 2) even though it was very prevalent, occurring between 4.0 and 12.0 times per ten minutes of talk. This was because the end of semester plenary session revealed that participants had said it to indicate their status, in the sense of “My opinion is in the category of ‘agree’” or “I am a member of the group of people who agree”.

Due to repeated exposures in the instructions and input, it was predicted that the input-then-task condition would have more instances of correct use of the form, ‘*agree with*’, and/or fewer instances of incorrect use, due to the repeated exposure.

When examining the transcripts, if an utterance contained the verb *agree* and no preposition was needed and none given, it was not counted. If a preposition was needed and supplied

correctly (with+NP, to+VP) it was coded ‘correct’, if it was needed and not supplied or needed and incorrectly supplied (to+NP, about+NP) it was counted as ‘incorrect’. No cases of the incorrect ‘agree+with+VP’ were found.

Table 6.9 has the results, standardised to the number per 10 min talk. The bottom row (GRAND TOTAL) indicates an apparent effect of input condition. Contrary to expectations, the task-then-input condition had 46 uses of AGREE+PREP, 30 of which (64%) were correctly formed, whereas the input-then-task condition had only 28 uses, 15 of which (52%) were correctly formed. This data, like L1 use, above, was not parametric, so Mann-Whitney U tests were conducted on the overall data. The null hypothesis held that input condition had no effect on: (1) the total number of the phrasal verb AGREE+PREP, (2) the number of correct/incorrectly formed versions, or (3) the percentage of correctly formed uses. Table 6.10 has the results of the Mann-Whitney tests, which do not allow me to reject the null hypothesis.

Table 6.9. Correct and incorrect instances of AGREE+PREP (per 10 min.)

Lesson	Topic	Task	Input-placement condition sequence						TOTAL
			Task-then-input			Input-then-task			
			A+P correct	A+P in-correct	Total (% correct)	A+P correct	A+P in-correct	Total (% correct)	
			Group 1			Group 2			
1	Single-sex education for girls	VC ₁ original	2.3	4.5	6.8 (34.4)	2.8	4.3	7.1 (40)	13.9 (36.7)
2		VC ₁ repeated	7.5	4.9	12.4 (60.4)	5.3	6.8	12.1 (44)	24.5 (52.2)
Subtotal 1: A+P use for VC ₁			9.8	9.4	19.2 (51.0)	8.1	11.1	19.2 (42.1)	38.4 (46.6)
			Group 2			Group 1			
3	Gender and language	VC ₂ original	5.3	1.5	6.9 (77.6)	5.4	2.3	7.7 (70)	14.6 (73.3)
4		VC ₂ repeated	14.8**	6.2	21* (70.2)	1.2**	0	1.2* (100)	22.2 (72.1)
Subtotal 2: A+P use for VC ₂			20.1	7.7	27.1 (74.2)	6.6	2.3	8.9 (74.2)	36.0 (74.2)
GRAND TOTAL			29.9	17.1	46.3 (64.6)	14.7	13.4	28.1 (52.3)	74.4 (59.9)

Note: items marked with an asterisk * or double asterisk indicate significant differences

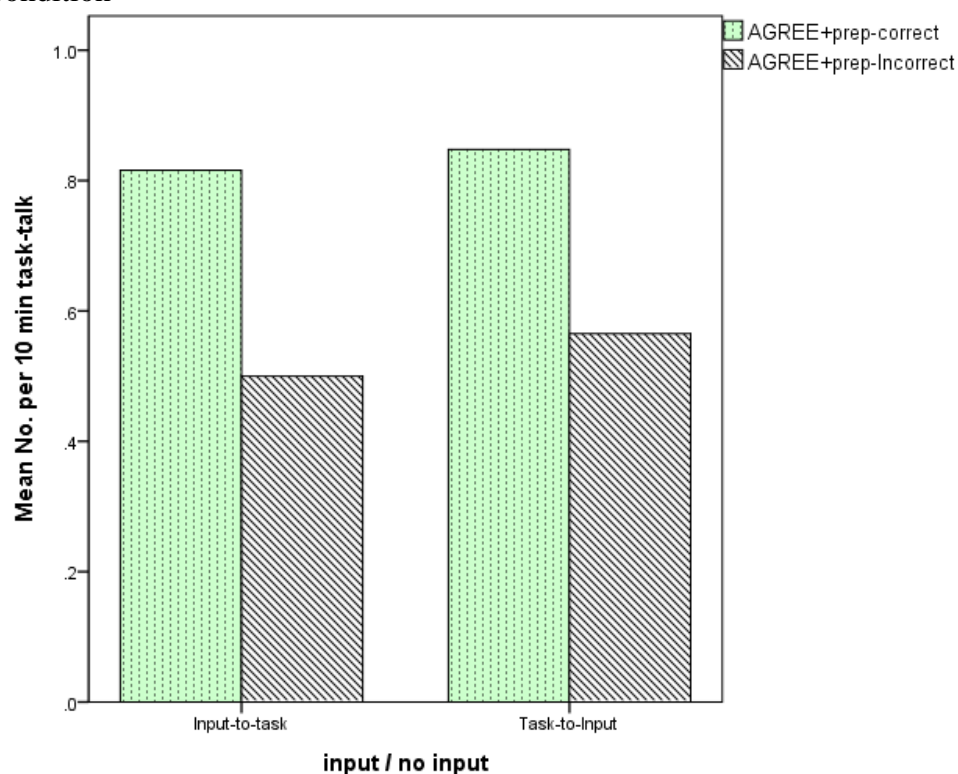
Table 6.10. Results of Mann-Whitney U tests investigating AGREE+PREP use across Input Condition (task-then-input vs. input-then-task)

Variable	Mann-Whitney U statistic	Exact Sig.
AGREE+PREP total	32.5	0.318
AGREE+PREP correct	32.0	0.383
AGREE+PREP incorrect	30.0	0.535
AGREE+PREP % correct	22.0	0.589

Looking across each row for each lesson from 1~4 in Table 6.9, the total number of uses seems consistent from one input condition to the other, except for the final lesson, Lesson 4 (VC₂ repeated task). This lesson, indicated with a single asterisk (*) in Table 6.9, had 21 uses in the task-then-input condition and only one in the input-then-task condition. This was found to be statistically significant (Mann-Whitney U= 5.5, p=0.011). In the same lesson, indicated by a single asterisk (**) in the table, AGREE+PREP was formed correctly 14 times in the task-then-input condition and only once in the input-then-task condition. This, too, was found to be a statistically significant difference (Mann-Whitney U= 8.5, p=0.038) favouring the task-then-input condition. It is not clear why task-then-input was preferred only in this one lesson, and it is also not clear why the participants in Lesson 4, input-then-task condition made so few attempts to use AGREE+PREP, particularly because in Lesson 2, VC₁'s repeated lesson, all the variables increased over Lesson 1 (VC₁ original).

Figure 6.3 shows the mean number of correctly and incorrectly formed AGREE+PREP phrasal verbs. It confirms the similarity of the effects of task-then-input and input-then-task conditions.

Figure 6.3. Correctly and incorrectly formed AGREE+PREP combinations per input condition



To sum up this sub-section, meeting the target form, '*agree with+NP*', repeatedly in pre-task input did not have a statistically significant effect on its correct use in the task. Qualitative evidence revealed that patterns of AGREE+PREP were highly variable within participants, and this may have overwhelmed any effect of input placement.

6.2.5 Summary of quantitative input-condition data

Overall, neither the input processing activity nor where it was placed had an effect on the output in a subsequent task. Input condition did not affect the number or type of LREs, the amount of L1, the use of targeted items from a language awareness activity, or the amount of correct/incorrect use of task natural grammar items, the '*agree+with+NP*' phrasal verb (a form of implicit learning). This lack of impact of the input processing activity requires explanation. I start by turning to the voices of the participants of the study to see how they experienced and used the input provided them.

6.3 A qualitative analysis of the role of pre-task input processing in task performance

One of the aims of this research was to investigate how learners *used* input in task performance in order to make recommendations for lesson/syllabus design. The results so far suggest that input played a very limited role in task performance. Next to no targeted items from the language awareness (LA) activities were used either in the original task condition or the repeated task. Because the pre-task input was not used as predicted, a major focus of the qualitative analysis is to determine why not. This section investigates why input was underutilized as a resource.

To do this, I will turn to the words of the students, from the end-of-study interviews and from their metacognitive awareness (MA) raising task. The interviews were done at the end of the semester. Participating students who volunteered to be interviewed were played two or three portions of their task performances. Participants were then asked questions according to a semi-structured interview model (Appendix 5) which allowed for talk to go off topic in order to explore emergent themes. Participants were also invited to ask the interviewer questions related to the course, the study or to their own learning. In all, ten participants volunteered. In contrast, the MA intervention was done during class time, so all participating students did it. The MA task had the benefit of requiring participants to discuss their thoughts with other participants rather than the instructor, allowing for triangulation with what was reported in the interviews. Transcripts of participant responses to interviews and the MA intervention were combined with video and audio transcripts of classroom activity when needed in order to explain the missing role of input in the study.

6.3.1 Student observations and perceptions of the input

One key question for the study concerned what students felt was the role of the input, and how they used it when doing tasks. From the end-of-study interviews, and the MA task performances, the following themes emerged:

1. Participants viewed the input as interesting because it was authentic.
2. Participants suggested making use of the language awareness (LA) items and transcripts in a more deliberate way.
3. Many participants did not prefer input-then-task or task-then-input order; most of those

who expressed a preference preferred to have the input first.

4. The input-processing activity did not impact on L1 use because L1 was used for a variety of reasons

6.3.1.1 Authenticity

The participants all reported they liked the input scripts, describing them as ‘*real, natural, native speaker, helpful*’. One participant, Lucky, called it “*complicated*” but she also said it was good to have.

Extract 6.2. Lucky: native speaker input is helpful and complicated

The reason why it is helpful is it is the conversation between the native speakers so we got to think that ah this is the word the native speakers really use in their real conversation and that is the also reason why it is complicated to us, because we only have the words known from the BOOKS or some kind of classes. But the words from that recording is from the real life. So that is also the reason why it is complicated

Here Lucky gives both sides of the argument for and against using authentic materials: they are real and in the form that they take outside the classroom, and this makes them complicated. By distinguishing the input text’s words from the words she was familiar with in English textbooks and classes she identified the main reason participants liked the input, it was new and different. However, this difference and newness may have been what prevented their use. Lucky’s report that she liked the input scripts and felt that they were helpful was a standard response across the participants in the study. Nada, had very similar ideas.

Extract 6.3. Nada: native speakers use language students don’t get exposed to

The recording itself uh that's helpful because when you guys discuss you're going to use the ...expressions you use in daily life, the things we're not really exposed to

She pointed out that students in this context do not get exposure to authentic language. Even from the media, which is typically scripted and edited, there is almost no chance for EFL students like her to be exposed to the um’s, interrupting, and “you know’s” found in this input (cf. J. Willis, 1996b).

Jon emphasized the reality of the expressions and suggested that the textbook, designed to improve academic listening, was artificial.

Extract 6.4. Jon: the input scripts were more realistic than textbooks

I like more your conversations and script rather than the text- the textbook because um, because some I think it is more in REAL conversation expression so I like this ...textbook is um artificial

The primary reason I used scripts of native speakers doing the task in the input processing activity was that students need to be exposed to natural speech in order to contend with it (J. Willis, 1996b, Ch. 6). All interviewed participants mentioned the reality of the input, with all the um's, false starts and incomplete sentences as being the reason they liked the input. Some participants, like Jon and Lucky, criticized the disconnection between regular textbooks and real language. However, participants almost never used the expressions either in their original task performance or in their repeated task performance as reported in section 6.2.3.

The most likely reason the items were not used is they may have been too new for the participants. The three instructors chose the language awareness items from the transcript based on what we felt the students would not know (see Appendix 10 for the full list). Using teacher intuition (cf. Prabhu, 1987; J. Willis & Willis, 2007) we may have effectively identified items the participants did not know, but this also meant that these items would be too new for the participants to use. A more sensitive vocabulary test, for example the Vocabulary Knowledge Scales (Paribakht & Wesche, 1993), may have uncovered smaller gains in lexical knowledge (e.g. increases in recognition knowledge) that the standard of 'use' is not sensitive enough to determine.

6.3.1.2 Obligatory use

To address the lack of use of the targeted items in the data, the students suggested that I require the use of these items. They felt that in order to learn the targeted items, I should have assigned an activity in which they had to be used, such as writing them in a new sentence, or made their use in the task mandatory.

This theme emerged from the talk in the metacognitive awareness (MA) activity. Students were asked about using the input. Question 4 (see Appendix 3) read: "*When you read the*

scripts and do the language awareness activities, do you try to remember new expressions, grammar or words?" Their MA task materials provided the question and a 4-point scale for the participants to indicate their level of agreement (1=No, 4=Yes) along with a place to make notes. Students indicated their level of agreement by circling their answer on the sheet and then talking about the question with their group members. I collected these survey sheets once the task was done. No participant circled 'YES' on their sheet. Although the MA survey was designed to elicit classroom talk and to implicitly raise awareness of the purpose of the task, and was not designed to be a reliable research instrument, the fact that no one indicated they tried to remember the targeted items is telling.

Extract 6.5 contains part of Mary's public report to the class summarizing her group's discussion of the MA questions. In this report, she expresses the ideas of her fellow students about the VC task and the LA items. She also implies her group's approach to language learning in general.

Extract 6.5. Mary's presentation of her group's opinion, 'learning is memorisation'

We think the teacher asks us to give our opinions about the topics, talk in English and makes us try to agree on one, new opinion, to make us debate actively so that we can improve our English speaking skills and practice to use the expressions and new words in conversation. We know the importance of memorizing the new words or expressions, but we usually don't spend enough time memorizing them. Because it's annoying job and we don't want to do it.

Mary and her group understood the dual intentions of the VC task: to encourage talk and to improve English. This report also demonstrated Mary's group knew the VC task was a chance to practice expressions and new words. She then equated learning the new words from the input with memorizing them and said they do not like to do it. This suggests that if learners feel they must memorise to learn, they might not try out new things without first memorising them. This 'memorise-learn-use' approach sharply contrasts with the 'use and learn' view underpinning TBLT.

Equating learning with memorizing was also a common theme for the participants in the stimulated recall interviews. They were asked if they tried to use words or expressions from the script or LA activity during the tasks or in the next lesson. In Extract 6.6, Jessie tells why she did not use the new items.

Extract 6.6. Jessie – testing beforehand will lead to use

Jessie

I think that when- when we learn that expression, and after that I- maybe discussion is one of the part of practicing that expression but may- I couldn't because that expression is not mine- I think that's not mine so

Researcher

How do you think you can make it yours?

Jessie

Oh it's very sad but maybe- test? ...I need some pressing about that you must speak and use that expression ... then I- ...Maybe that discussion test you know we always have that test. That test will be good for using that new expression. And after that we we we have ah discussion again then we can use more that

Jessie, like the members in Mary's group, realised that the discussion phase of the VC task was an opportunity to use new words and expressions. However, she felt she could not make the most of that opportunity because the words were not 'hers'. She asked for more emphasis on using those targeted items: "*I need some pressing about that you must speak and use that expression*". She needed the outside pressure of a test in order to remember the expressions, make them hers and be able to use them. Jessie felt that testing would drive her learning. She admitted, "*maybe discussion [the VC task] is part of practicing that expression*", but her comments suggest the discussion should be after the test. In addition, her choice of the word 'practicing' and not 'learning' is telling. Like the members of Mary's group, she appears to have equated learning with memorising and speaking with practicing. She also suggested repeating the task after the test so they could "*use more that [the target items]*".

Late in the semester, during the Lesson 6 task performance, Jessie told a story of learning from their speaking test. In it, she asked if the others remembered her problem remembering the word 'expert' during their last speaking test (a group discussion format). This story is in Extract 6.7 below. It reflects Jessie's belief from Extract 6.6, above, that she learned well when pressured. The item learnt was the word 'expert', a word she probably had some, incomplete knowledge of before the test, given Lucky and Cathy's memory of Jessie saying 'experee' (turns 145 and 146).

Extract 6.7. Jessie's story of learning 'expert' during the test

Turn	Speaker	Utterance
144	JESSIE	did you remember ? that I said in discussion test expert- expert and (S? ah !) and you guys EXPERT ! I'm I very sensation (laughter) I get very
145	LUCKY	no you were like ex-//experee\\ what is ? we are like
146	CATHY	//experee\\
147	JESSIE	so you guys said EXTRA what what ? example
148	CATHY	and they you said L1 <i>jeonmunga</i> 전문가 (tr=expert) we are no no it's just //EXPERT \\
149	LUCKY	//EXPERT\\
150	JESSIE	(laughs) so that night I go to my- I went to my mother and I said mother // do you \\
151	LUCKY	//mother\\ (hh)
152	JESSIE	know L1 <i>jeonmunga</i> 전문가 (tr=expert) in English ? mother said expert (laughter) ding
153	LUCKY	(laughs) sensational

Other interviewees echoed Jessie's call for an explicit requirement to use the targeted expressions from the input. Nada, for example suggested reviewing them the next lesson by having students choose some of the items and write new sentences with them.

Extract 6.8. Nada: writing new expressions into sentences

In language awareness we got a seven to eight unfamiliar expressions ... maybe we can pick up three of them and make a sentence [in the next lesson] ...

Nada's formula for learning from input, therefore, also included some deliberate activity, but not memorising, but producing them in the slower, more deliberate context of writing. For her, learning the new item meant practising it by herself on paper before using it in interaction.

These comments were made despite the fact that the course textbook also included vocabulary items that were used in classroom activities and were tested. Yet, the participants expressed a need for more deliberate work on language based on the VC task input. This was perhaps due to their belief that the native speaker input was more 'real' than the textbook. However, for instructors who practice task-based teaching, this is a timely reminder not to ignore the 'Deliberate Study' strand of language teaching (Nation, 2007, 2009) as, for among other reasons, this is still held high among student expectations, and may enable the transition from input to output. These participants certainly felt they needed practice time, alone – either memorising (Jessie) or writing (Nada), prior to using new items publicly. The students' voices in this section challenge the underlying assumptions of some cognitive views of TBLT that hold that noticing leads to learning and use. Jessie and Nada's views indicate that more

explicit forms of vocabulary instruction than consciousness raising (Rutherford, 1987) should be integrated with TBLT. They also suggest that students' expectations and preferences should be included in instructional decisions, and this moves the discussion of syllabus planning to a local, ecologically valid level.

6.3.1.3 Input order

Participants were asked if they preferred the input-then-task order or the task-then-input order. They generally preferred input first, but some preferred task first and three of the ten participants did not notice the change in order at all. Jessie is a typical example of those who preferred input first.

Extract 6.9. Jessie's preference for input-then-task

I think first script and after discussion is better ... because, uh we we can learn new part of expression and we we think about it before we discussion ... Then I can talk about it. I can talk about that and using that word, I will using that word but that's not easy but I can think more

Jessie claimed that if the input were first she would use the new words that were given before. However, in the data set, she never did this. She did say it was not easy to use the new words, but she liked the idea of having the input first, so that she could potentially use them. In Extract 6.6, above, Jessie said she had not made the new items her own so she could not use them. Here, she claimed that input-then-task was preferred so that she could. So, the immediate use of IPA-targeted items were: the lack of explicit requirements to use them (Extracts 6.6 and 6.8), the lack of additional activities in the language awareness (LA) that require their use, and, as will be discussed below, the lack of a model of mining from input and course materials.

In contrast, Janine preferred to have the input second. In addition, she expressed doubts as to the usefulness of the language in the scripts, another limitation on learners using the items.

Extract 6.10. Janine on pretask input as source of ideas, not language

I don't know it's [the input] helpful [for learning English], it's interesting, it's a view from- view of the foreigners so different from Korean. So uh rather than helping improving my English skill it's just very interesting

Janine claimed that the opinions expressed in the input were interesting to her because they were often different from what she was used to (Korean opinions), but she felt the input was not going to help improve her English, and that it was just interesting. Participants with this view of the input will not try to use the items targeted by the instructor or try to notice other items in the input.

Lucky raised the issue of forgetting across lessons. She told the members of her MA intervention team that she could not remember the previous lesson until the instructor reminded her.

Extract 6.11. Lucky cannot remember the previous lesson

I can't remember exactly what I learned last class, so when Shaun [the teacher] reminds us beginning of the class I start to remember

This may explain why the repeated tasks also had no evidence of use of the targeted items given in the previous lesson; the students had not been required to memorise or use them, and simply forgot.

So, preference for input before the task was indicated by participants who felt they could use the language from the input-processing activity, even if they did not actually use it.

Preference for input after the task was expressed by participants who felt the input was interesting from an information-content perspective but not useful linguistically. Others did not notice the order of input and some needed help to remember the topic. These differences in alignment with the input (Atkinson et al., 2007) may account for its overall lack of impact.

6.3.1.4 L1 use

It was predicted that L1 use would decrease in the input-then-task condition because participants would have samples of L2 available to use. However, results above demonstrated that learners did not use the language from the input in their task performance. Therefore, participants relied exclusively on their existing linguistic resources when doing the task. This left the participants resorting to L1 to fill in gaps.

The participants discussed the use of their L1 (Korean) both in the end-of-study interviews and in the metacognitive awareness (MA) task. Responses regarding the use of L1 indicated that participants felt they used Korean only when they needed to compensate for a lack of L2

English. This was expressed by Sunny and Winnie in Extracts 6.12 and 6.13. Both Sunny and Winnie are in the team that used the most L1, A3. Unfortunately, none of this team's members was able to participate in a stimulated recall interview; however, their metacognitive awareness (MA) intervention talk provided an idea of their approach to L1. Sunny referred to not knowing individual words and their meanings and Winnie pointed to a general lack of English proficiency.

Extract 6.12 Sunny used L1 as a supplement

[We use Korean] when we don't know about the verb or the textbook uh word's meaning to explain it ...

Extract 6.13. Winnie claimed a lack of English leads to L1 use

uh I think I always use Korean because I can't speak English very well so I don't know some expressions in English

Neither Sunny nor Winnie referred to their extensive off-topic L1 conversations in the MA task. However, their comments do account for on-topic L1 use.

The source of the lack of L2 was not mentioned by either of Sunny or Winnie, but several themes emerged from other participants who reported using on-topic L1 when a lack of L2 was created by speaking about an unfamiliar topic or by discussing it in more detail than they were used to. They also referred to social pressure as a driving force for using off-topic L1 when they could have used L2.

Topic unfamiliarity

Participants described a lack of familiarity with the topic as a problem for the use of English. At such times, they did not have enough vocabulary to be able to talk about it. The topics, 'Benefits of same sex education for girls', 'Gender and language', and 'Dangers of the mass media' (Sanabria, 2004), were complex social issues which many first year university students would not have ever discussed in detail, even in because of the Korean education system's heavy emphasis on memorization and test preparation (J.-K. Park, 2009; Song, 2012, and see Chapter 1). Jon, for example, in Extract 6.14 gives this as his reason for using L1.

Extract 6.14. Jon talking about the relation of topic familiarity and L1 use

If we face the problem, the situation that we didn't talk about, that moment we always uses- use the Korean, because we don't know the English word for that so that's point- that point, we always say in Korean

Jon indirectly raises the relationship between the topics, which he called “*the situation*,” and vocabulary. The input would have increased topic familiarity along with providing sample vocabulary. However, Jon admitted that they used Korean (L1) rather than look for L2 alternatives from the input.

Pushed output

The second reason for L1 use was participants pushed the topic of their discussion beyond the level of L2 vocabulary. Participants also reported that the VC task forced them to think more deeply about the topic and explain in detail. The VC task required an expression of opinion, justification of that opinion and then negotiation to reach a consensus. Due to their educational background (Chapter 1), the last two elements are processes they would not have been very familiar with by the time of the study. Both Lucky and Jessie referred to the depth of their discussions in their interviews, shown in Extracts 6.15 and 6.16. They were in the same team and both separately raised the issue of the depth of their conversation.

Extract 6.15. Lucky on the process of task talk in her group

Every time when I doing what I doing this task what I feel is like, what I feel and what I think is that, our group 5 is going too much deep. Like, mm uh going deeper means like we kind of touch the government things like policy- government policy and I think we are going too much deep and deeper.

Here, Lucky described the depth of the conversation and seems to be complaining about it. Whereas, Jessie, in her interview a week later, explained why the depth of the conversation was a problem, she could not continue it in English (Extract 6.16).

Extract 6.16. Jessie on the depth of the conversation prompting L1 use

Actually speaking Korean that topic is not difficult. But I always think about more DEEP that when I speak in English, so it's hard to express my opinion.

The deep conversation was a goal of the task, so in this sense, it was positive to hear this is what they had done. In addition, looking at Lucky's opinion in Extract 6.9, we see that she used the same sort of process to give this opinion as she would have done in the VC task; she gave an opinion in the first statement, and then clarified what she meant with the second, indicating that she had developed a habit of clarifying her ideas. This direction of the discussion, however, would not necessarily be congruent to that of the native speakers, meaning that vocabulary from the transcript of their talk might not be helpful.

The vocabulary from the script would have been maximally helpful to participants who agreed with one (or more) of the native speakers on the script for the same reasons they gave. In such cases, the participants could have copied what was said verbatim. However, this copying never happened, most likely because the participants genuinely wanted to give their own opinions. For instance, Sunny, when doing the MA task, reported she never simply agreed with others to help the task go along (Extract 6.17).

Extract 6.17. Sunny always wanted to give her opinion

I can't agree with someone's opinion, it's different ... so we have to make a sentence to strongly agree...opinions are different but we have to DO our task so I have to talk

Taken with Extracts 6.15 and 6.16, Sunny's comment demonstrates that the learners aligned with the opinion gap element of the task and to talking (Atkinson et al., 2007; Atkinson, 2002, 2010). Sometimes, their ideas stretched beyond what they could say in English (Extract 6.17) prompting L1 use.

Social pressure

Another reason for L1 use, and for off-topic L1 use, was social relationships within the teams. The first type of pressure was to conform to their peers who share the L1. For example, in her MA intervention talk, Jessie stated that the social pressure of the situation, not wanting to appear strange, made her speak Korean (Extract 6.18).

Extract 6.18. Jessie – L1 use to fit in.

I think situation makes me ... speaking in Korean, if you and you and you speak in Korean and only I speak in English, I can be a strange person ... so I follow the situation ...

Jessie claimed that negative peer pressure inhibited her. However, her team (A5) actually used very little L1 in the VC tasks (see Table 6.8), meaning this situation would almost never have arisen. It is possible that she was referring to cases when the students were doing other tasks in other groups, or, perhaps she was referring to more generally to her learning-history, or this itself is an attempt to fit in with her (new) group members during the MA task by not admitting that her group rarely used Korean.

Nada was a member of a high L1-use team (A2). When asked why she used Korean, she claimed L1-use was based on communicative pressure and their temperament. Extract 6.19 is how she explained this to me.

Extract 6.19. Nada – using L1 to ‘hang out’ and to ‘complete thoughts quickly’

Nada

You know Michelle, Dave and Will got a strong character, we use to ...hmm we used to mess up and tried to hanging out with the group member not discussing the main topic. While doing that we got some expression in Korean that we can't do in English or you know Michelle and I, how can I say that, we can't wait a second to come up with the English word,

Researcher

So you're impatient?

Nada

Yeah, we want the word coming right now.

The first part of her comment refers to the off-topic social talk her team used. Transcripts of this group revealed that, unlike Jessie and Lucky's team, they often ‘followed the leader’, that is, when one person started in L1, the others followed in an effort, as Nada put it, to “*hang out*”. She later mentioned that she wanted the words to come immediately, and could not wait or spend the time and mental energy to recover them in L2. Doing the task and any off-topic talk in English would help develop her automaticity, or as she put it “*maybe we can speak more fluently*” (Nada, MA task talk), but she did not use the learning opportunities presented by the task to do this due to social needs at the time.

Another member of the same team, Michelle, echoed Nada's reasons for using L1. Michelle, in Extract 6.20, said that she used L1 a lot when her friends were in the same group.

Extract 6.20. Michelle – Korean appears immediately when talking to friends

*So often especially when my major friends are in the same group. Korean just comes up
POP POP POP*

Nada is one of Michelle's 'major friends' (they study the same major) and they are also in the same VC team, as with the case of Lucky and Jessie, Michelle and Nada corroborate each other's statements.

Summary

The VC task was designed to have participants discuss topics from their course book and express opinions, justify their opinions and negotiate a consensus response. For the most part, this was done in L2, but there were times when L1 emerged. It had been predicted there would be less L1 used by groups that had received the input-processing activity before doing the task. However, the results did not match predictions. Part of this was due to the participants' reasons for using L1. The unfamiliar topics and depth of discussion both prompted L1 rather than pushing the participants to develop their L2. In addition, using L1 was the participants' way of fitting in socially.

As to why the participants did not use targeted expressions from the input there are a few explanations. First, participants did not view the items as something they needed to learn. They expected the instructor to use more direct approaches to push learning and use. The items were new and the participants suggested more 'silent time' with them in order to be able to use them. In addition, the participants' opinions may have lacked congruence with those of the native speakers on the script so some items may not have been relevant to the meanings the participants wanted to express. However, many of the targeted items were 'topic mobile', that is, they could be used in other contexts (e.g. '*beyond/within the scope/range of*'). Finally, learners may not have used the scripts once the input-processing activity ended. This point will be confirmed in the next section.

6.3.2 Classroom observations

Having listened to the participating students' voices about the input and its role in their learning, I felt that examining what they actually did with it during their VC talk might be helpful, in order to determine why so little input was used. During the initial data analysis, it

was felt that the participants' physical actions (e.g. reading or writing) during the task may have influenced their ability to make use of the input. Most importantly, for learners to use input, they would either need the expressions visible and accessible during the talk or have strong recall capacity for the item. As the latter would not be possible for a newly learnt item, accessibility during the task is of importance. To investigate whether or not learners referred to the targeted items, which were on both the input-processing activity sheets and on the whiteboard, I reviewed the videos of each lesson.

During the data collection, one group was selected at random and videotaped per lesson. In subsequent lessons, the videotaped group would be changed so that all groups were video recorded at least once during the study. The videos were not ideal, in that often one of the group members' face or desk was not visible, and determining what participants were looking at was sometimes problematic as I had only one camera pointing at one group and no other video equipment, such as a digital tabletop (Seedhouse & Almutairi, 2009).

Still, it was usually possible to determine what the participants were holding, writing on or looking at. The VC task prompts (Chapter 4) were on a single sheet of paper and had a distinct bold outline in the shape of a rectangle with rounded corners that could be easily identified. The input-processing activity sheets were separate from the VC prompts, were three pages long and consisted of a long transcript (See Appendix) with the LA questions and answers on the last sheet. When the participants held them, it was clear there were multiple pages. Figures 6.4 to 6.12 show Team A3 and Team D1's activity during the input (including LA) phase and during the VC task. They illustrate the typical actions of the group members, and, the effect these actions ultimately had on target item use.

Figure 6.4 shows Winnie doing the LA activity. She is holding three sheets of paper stapled together. This is the input transcript along with the LA questions. Sunny has her back to the camera, and Nancy, sitting to Winnie's right, is hidden by Sunny.

Figure 6.4. Winnie holding the input sheets containing transcripts during pretask input



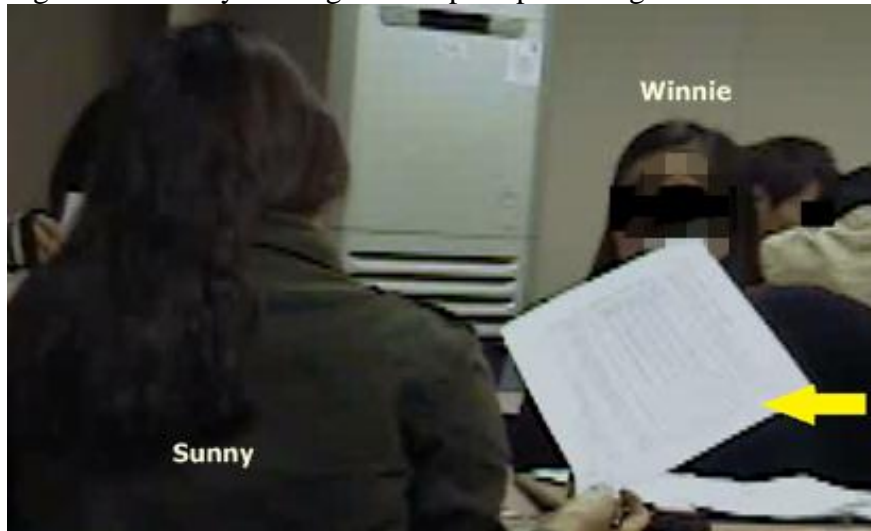
After the LA has been done, the answers were corrected by the instructor in a teacher-to-whole class format. The participants looked toward the instructor who was standing in front of the white board, and copied down any answers they did not get correct. All LA answers were written on the board and left there until the class was dismissed. In this lesson, the board is to the right, and behind the camera. Figure 6.5 shows the positions we would expect to see if the participants turned to look at the board during the task.

Figure 6.5. Teams A3 (front) and A4 (rear) facing the board during feedback on LA



Figure 6.6 shows Sunny holding the VC task prompts in her hand, indicated by the arrow. We cannot see if she is looking at her copies of the input sheets on her desk.

Figure 6.6. Sunny holding the VC prompts during VC task talk



We can see in Figure 6.7, that Sunny was the one who wrote the final versions onto their answer sheet. Although we cannot see her paper in the picture, her pencil is circled. We cannot see from these figures, but it was evident in the video, that Winnie had placed her prompts sheet on top of the other sheets and did not move it. So, in Figure 6.7 we can conclude that she is looking at the prompts and not the input.

Figure 6.7. Sunny writing the answers



The series of stills from Team A3 are somewhat inconclusive due to body and camera positions, but we cannot see the members of the group ever looking at the board, or looking at the stapled sheets of the pretask input during their VC task. It is possible that Nancy had

them open or that Sunny looked at them briefly. I could see from the video that she never lifted them or folded them over, meaning, if she did look; it was at only one page.

The next series of screen shots show Team D1 during Lesson 11.2. Here the camera is on the side ledge of the classroom and the white board to the left and behind the camera. The angle was a little better in this room so that the camera could look down onto their desk, and no one has their back to the camera.

Figure 6.8 shows the three members of the team looking at the multi-page input. This is during the LA portion of the pre-task input processing activity. Eunice has input laid out on the desk opened, so it has been circled to indicate the fold. Monty is flipping through multiple sheets of paper stapled together. Likewise it is clear that Mary is looking at a few sheets attached together. This is the input transcript. The input and task prompt can be differentiated from each other on the video in this recording too, based on staples and folds and the darker rectangle with rounded corners of the prompt.

Figure 6.8. Team D1, looking at the pretask input transcript and LA



Later, when the instructor is giving feedback to their answers for the LA activities, we see the class look toward the board to the left of the camera (Figure 6.9). This body position should be approximated if a participant were to look at the board for the new vocabulary.

Figure 6.9. Teams D1 and D2 looking at the whiteboard during LA feedback



It can be seen in Figures 6.10 ~ 6.12 that Team D1's task performance visibly centres on the VC prompt sheet and not on the pre-task input.

Figure 6.10. Team D1, VC task talk – 'Mary: *"I have experienced becoming stupid"*'



Figure 6.10 shows the team during a humorous moment in their talk, about the prompt:

"Young people are becoming more and more stupid because they watch TV too much and they don't know how to use the information they get." Mary's reason for agreeing with the prompt is that she feels she has become more stupid, saying *"I have experienced becoming stupid."*

The upward arrow points to the answer sheet, on which Eunice was responsible for writing their final revised version of the prompt. The video analysis revealed she placed the answer sheet, essentially a blank ½ sheet of paper with lines on it, on top of the input scripts. We can see the folded and stapled corner of the input in the circle. The arrow pointing left points at Eunice's sheet of prompts. Mary and Monty have both placed their prompts directly on top of their input.

Figure 6.11. Mary suggests deleting ‘more and more stupid’



In Figure 6.11, we see the group trying to decide on a new version of the prompt. Mary has suggested deleting part of it, “*becoming more and more stupid*,” as part of their final version. As she said this, she moved her pen and scratched it out on the prompt. Monty is pointing at the prompt with his pen, and Eunice’s gaze is at her prompt sheet that is on the left side of her desk.

Figure 6.12 shows Eunice writing on the answer sheet while Mary is telling her a potential answer. She said, “*young people watch too much TV*”, and Eunice, without looking up, interrupted her and said “*too much TV and internet, okay?*” prompting “*okay*” from both Monty and Mary. The collaboration on the final version of their revision to the prompt took place with similar body positions as those in Figures 6.4 – 6.7. They look at each other, particularly early in the discussion. Then their gaze typically moved downwards to the prompt as it was written. They focus on the prompt, examining its language more closely and make suggestions about changes to it while focusing their gaze on the sheet containing the prompt and on the answer sheet where Eunice was writing. At the end of the VC task, when all five prompts had been done, Eunice gave the sheet to Mary for the presentation. This confirmed that it was the answer sheet that Eunice had been writing on, and that this sheet had been covering her copy of the input.

Figure 6.12. Mary dictating while Eunice writes and clarifies



While the video data is not systematic, reviewing the videos of the VC tasks showed that one of the main actions that inhibited the use of the new items was that the participants simply did not refer to the input while they did the task. There were no cases on video of participants referring to their copy of the native speaker transcript and looking for a word or expression to use. Nor were there cases of participants looking at the whiteboard and then using one of the answers.

The input that was used was the words and grammar structures of the prompt in their VC task talk (see Chapter 5). We see from these stills taken from these two groups' pre-task input processing activity and VC task talk why this is the case. The students have the prompts in front of them the entire time, they are focused on fixing the prompt, and they never refer to the transcript. It appears that for students to use the targeted items: either the answers to the LA activity should have been included in the prompts, or, space for writing the answers should have been provided on the same page as the prompts. As it was, participants relied exclusively on memory, i.e. their current language resources, while doing the VC task. The presence of input played no role in task performance, whether pre or post task.

6.3.3 Summary of the qualitative input-condition data

To sum up the qualitative data regarding the lack of input use in the study, it was found that the participants, both in their interviews and in their MA tasks, felt the input was interesting and helpful. Some mentioned they could use words and expressions from it later, but there

was only a single case of them doing so. Participants suggested ways to make the targeted input more relevant to their talk: by either requiring participants to use the language awareness (LA) answers in review activities at the start of the next lesson, or by requiring these items to be produced on a test. Participants also said that such requirements to memorise or learn new things would be boring, tedious or difficult. Some participating students expressed preference for having the input first, ostensibly for being able to use the input, despite the fact that they did not. Others preferred to have the input after doing the task because they felt the native speaker opinions were interesting but the language in the scripts was not useful for them. This different perspective on the input was an unforeseen factor mediating the use of the target items.

A closer examination of the physical movements of the participants and visual reference to the input activity resource used in the study revealed a potential reason for the non-use of target items – the participants simply did not refer to the input during their tasks. Their potential affordance as an aid in task performance went unrealised. The video analysis suggested that placing the target items into the task prompts (which Chapter 5 reported were heavily mined), or physically closer to the prompts (i.e. on the same sheet/page) will be needed in order to get students to use these items in their task performances. The other option, suggested by the participants, is to explicitly require their use, but this diminishes from the spontaneous, natural speech nature of tasks, and works against the tenets of TBLT.

6.4 Discussion of the input-placement manipulation

When participants were given pre-task input, it was expected that they would use the native speaker input as a guide for their task performance. The participants would know what was expected, and therefore they would have more cognitive resources available for focus on form leading, among other things, to an increase in language related episodes (LREs) (Kim & McDonough, 2011). In addition, the input-processing activity was based on a sample performance of the same task. Therefore, the participants had access to task-related language and task procedures, and it was predicted that they would use less L1 in the input-then-task condition.

No statistically significant difference in the presence of LREs across input placement

conditions was found. This was unexpected when compared to Kim and McDonough (2011). In their study, they found that pre-task modelling increased the number of LREs. To explain the different findings, we should turn to the nature of the participants, the input and the VC task.

6.4.1 Participants' previous experience with the tasks

The first factor that may explain why the results did not match predictions was the previous experience of the participants with the tasks. Kim and McDonough (2011) worked with middle school students, who had four years of school-based English learning and some extracurricular classes but in the classroom were given “few opportunities to communicate using spoken or written English” (p.187). The participants in the present study were first year university students – who had six additional years of the same type of instruction plus one semester of university instruction. They were thus more experienced language learners. From the first day of lessons in the course they were taking, they had done opinion and information gap activities. It has been reported that Korean students in general have little chance to interact in English during their pre-university school years (Jeon, 2009; S. Park & Manning, 2012; Shin, 2012). Moreover, although Kim and McDonough (2011) did not mention if the students had done dictogloss, information gap and decision making tasks before, given the location of their study, it seems plausible that they had not.

If their students had had no previous experience and these results were obtained, then experience with the task may be an intervening variable influencing the model's impact. The model may serve as a vicarious experience for learners who have not done that task before; however, learners with experience doing the task may not benefit from the pre-task model as they have their own internal model. The participants in my study all fell into the ‘experienced’ category as they had all been trained to do the VC task, twice, prior to the beginning of the study. So if the model's benefit is as a procedural example, the facilitative effect of modelling was eliminated in my study. Because the training lessons eliminated any confusion among the participants about what was expected in the task, both groups, the input-then-task and task-then-input conditions, had nearly equal resources available to them during the task performance and there was no need to deploy resources anywhere except to the task. It is this that may account for the similarity in number of LREs and amount of L1 across the two conditions. Some learners did not even notice that the placement of the input had been changed, implying that the tasks were experienced equally regardless of when the input

occurred.

6.4.2 Input-related factors

6.4.2.1 Modelling LREs in input

The nature of the input may have had an impact on the lack of helpfulness of the models. Kim and McDonough's (2011) models were not models of spontaneous task performance as called for by Willis (J. Willis & Willis, 2007; J. Willis, 1996a, 1996b), rather they were scripted by the researchers to ensure the models showed successful resolution of LREs. The participants in Kim and McDonough (2011) who viewed the input had more LREs. I did not write a script of the input before recording it, but rather I transcribed *post hoc* a recording of three native speakers doing the task spontaneously. These scripts did not involve any LREs because, as one would expect, there were no language breakdowns among the three native speakers. In this study, where there were no LREs modelled and just an example of the task performance was modelled, the number of LREs was not changed. Hence, the increased number of LREs among students who received pre-task models of LREs in Kim and McDonough (2011) was perhaps due to the modelling of the process of doing LREs, and not due to a reduction in cognitive load caused by the model freeing up resources for attention to language.

6.4.2.2 Factors preventing input being used

In this study participants read the transcript, analysed it for target items, listened to the instructor explain the correct items, and then drill some of these items as recommended by Willis (1996b). Yet the items were not used. Some reasons were suggested.

Physical placement

The physical location of the target items after the LA activity was finished prevented their use. Because they were on a different sheet as the prompts, the learners never looked at them. They covered the input sheets or put them under the desk, making referring to the items impossible. They were never seen on video looking up at the whiteboard for an item, either. As none of the items were task essential (Loschky & Bley-Vroman, 1993), the target items could be avoided during task performance.

The language awareness activity comprised of a consciousness raising (CR) activity followed by some controlled practice was not enough to overcome the 'newness' of the items and

enable them to be recalled. And, without having them in sight, participants opted to use items they already knew.

A way to encourage new item use might be to move the answer sheet to the same piece of paper as the VC task, perhaps leaving blanks on the page under the prompts, for the students to write the final answers on. If the targeted items remain visible to the learners as they perform the task, there might more use of them in the task. This small adjustment could be easily tested empirically.

Mental energy expended for learning

One point not addressed in this study is that when participants did not get the item correct in the LA activity, and just got the answer from the instructor they may have expended very little mental energy determining the correct answer. To address this, the correct item could be re-emphasised by having them write a sentence, as suggested by the participant Nada (Extract 6.8). This might well lead to both immediate and long-term benefits for the learners. This issue could be tested experimentally in future studies.

Requiring the use of the items in a test, as Jessie suggested, would be helpful for long term learning, because it is a form of deliberate study (Nation, 2007, 2009). It would, however, change the learning process of the item from ‘in-task’ learning to ‘out-of-class’ learning. Tasks are intended to promote learning in and of themselves; and items used in tasks should therefore be available to the learner on a test. Jessie’s suggestion to require students to separately memorise items for a test would then circumvent one of the purposes of the task – to help learn the targeted items during the task cycle and have it available and automatised for later use (e.g. on a test).

Temporal distance

A third and final factor that may have contributed to the lack of use of targeted input was the amount of time between original and repeated tasks. A limitation of the study is that the repeated task was performed at the start of the following lesson, a few days later, unlike Hawkes (2011) who repeated a task immediately after an IPA. This was a necessity due to the constraints of doing research on existing classes who were studying a set curriculum. For participants in a ‘task-then-input’ condition, this delay may have been too long them to remember and use what they noticed during the post-task LA activities. So, Willis’ framework

which called for a post-task focus on form has not been directly assessed (J. Willis & Willis, 2007; J. Willis, 1996a, 1996b) in the context of a values clarification task.

6.4.2.3 Incidental acquisition of a repeated lexico-grammatical structure

The study predicted that participants' use of AGREE+PREP would have become more accurate over the course of the study. The item appeared in all the input scripts at least six times. Yet, there was no observed improvement. The explanation for this lies in the prevalence of incorrect counterexamples from other participants. For example, one participant, Sara, said:

Extract 6.21. Sara using the verb 'agree'

"I strongly agree number¹ one: number three and four but uh I'm not agree² with³ number two"

Extract 6.21 shows one sentence that contained: (a) a use of 'agree' without the necessary preposition 'with', at (¹); (b) incorrect use of the verb 'be' to indicate a state of agreement at (²); and, (c) the correct use of 'agree with' at (³). The range of Sara's use of the form means that the other members of her group, who did not correct her, were exposed to two incorrect and one correct use of the item. Relying only on implicit learning of this form was perhaps a mistake, when it could have easily been addressed in the LA activities or through immediate corrective feedback by the instructor. Although it was addressed with the participants at the end of the study, they had lost many opportunities for correct use of the item.

Prabhu (1987) long ago argued against the use of small groups and pairs because students would be exposed to poor language models, and it appears this may have happened in this case. To address this, other approaches to using the values clarification task to aid in acquisition may be helpful. It is to these we now turn.

6.5 Chapter summary

This chapter reported the results of one of the interventions of this study. I set out to investigate the effect of using recordings of native speakers doing the VC task on the way

participants performed one. Against predictions, there was no impact of the input processing activity (IPA), whether it was pre-task, or post-task. Upon examination of qualitative data, several reasons for the lack of effectiveness of input were proposed. These included:

1. The participants' perceived the input as ideational content and not as a language resource. They treated it as something they needed to understand and not as useful language.
2. The items were too new for immediate use and the participants expected the instructor to test the items or otherwise force them to be used in class.
3. The input was fundamentally different from other models; it did not model the in-task behaviour that was measured, i.e. LREs.
4. The words and expressions in the VC task prompt were mined extensively, similar to mining reported in other studies because they were task-essential. Use of targeted items from the language awareness part of the input-processing activity was absent because it was task-natural (Loschky & Bley-Vroman, 1993).
5. Use of L1 was idiosyncratic to teams and unrelated to input.
6. Students did not refer to the input transcripts while doing the VC task.
7. The correct models of AGREE+PREP in the input were insufficient relative to the number of incorrect models received from their peers.

This chapter has shown there was no effect of the input-processing activity on VC task performance. There were three other interventions: task repetition, procedural repetition and the performance of a metacognitive awareness (MA) raising task. The next chapter reports the effects of task and procedural repetition.

CHAPTER SEVEN: THE EFFECT OF TWO TYPES OF TASK REPETITION ON TASK PERFORMANCE

7.1 Introduction

This is the third of four chapters that make up the results and discussion. The previous chapter addressed the effects of an input-processing activity on task performance. It reported little or no impact of the input processing activity on task performance whether it was done pre- or post-task.

This chapter addresses the impact of two types of repetition: task repetition and procedural repetition (Kim, 2013), on task performance. First, it briefly restates the research design and related research questions. As in the previous chapter, the chapter presents quantitative answers to the research questions, followed by qualitative data to shed light on the results. It then discusses the findings with respect to other studies and theoretical perspectives and then concludes with a summary of the main findings.

7.1.1 Review of the research design

To remind the reader, the study involved three values clarification (VC) tasks, each done twice, across six lessons in an ecologically valid context. In the previous chapter, the first four lessons were investigated for the effect of an input-processing activity. The last two lessons involved a metacognitive awareness (MA) raising activity that was implemented to improve task performance. These will be examined in the next chapter.

Similar to Kim's (2013) study, two types of task repetition were built into the design: procedural repetition and task repetition. Procedural repetition involved repeating the same task procedure on a new topic. In this study, these were called 'original' tasks. Task repetition was defined as doing the VC task a second time on a given topic with minor changes made to the prompts – i.e. the procedure *and* the topic were both held constant. In this study, these were called 'repeated' tasks. Both procedural and task repetition conditions were repeated three times each through the study. Table 7.1 shows the arrangement of topics and tasks. Unlike Kim's (2013) study, which had a between-subjects design, one group doing only procedural repetition and the other only task repetition, this study is a within-subjects

repeated measures design. The participants did all six lessons.

Table 7.1. Arrangement of topic and task repetition type

Lesson	Topic	Task
1	Single-sex education for girls	VC ₁ original
2		VC ₁ repeated
3	Gender and language	VC ₂ original
4		VC ₂ repeated
5	Dangers of the Mass Media	VC ₃ original
6		VC ₃ repeated

7.1.2 Research questions

The main research questions addressed in the chapter are as follows.

1. What is the impact of procedural repetition (Kim 2013) on VC task performance?
 - a. Does procedural repetition increase the number of LREs?
 - b. Does procedural repetition decrease the amount of L1 used?
 - c. Does procedural repetition increase the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?
2. What is the impact of task repetition (Kim 2013) on the task performance?
 - a. Does task repetition increase the number of LREs?
 - b. Does task repetition decrease the amount of L1 used?
 - c. Does task repetition increase the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?

Because it was found in the previous chapter that the students did not use targeted items from the input processing activities, this will not be addressed in this chapter.

7.2 Quantitative results and analysis

7.2.1 The effect of procedural repetition

Three VC tasks were done twice (six in total) over the length of the study. The last task (VC₃) was preceded by the metacognitive awareness (MA) intervention, which will be discussed in

detail in the next chapter. It will be shown in the next chapter that, unlike the input-placement intervention, the MA intervention did affect task performance. This means that VC₁ and VC₂ (Lessons 1-4) investigate procedural and task repetition while VC₃ (Lessons 5 and 6) can only investigate task repetition (with MA condition as a second independent variable) because the procedural conditions were different from those of VC₁ and VC₂. The results are in Table 7.2, on the next page, with the MA intervention data shaded darker.

Table 7.2 shows a decreasing number of LREs (per 10 min of VC task talk) from lessons 1-4, an increasing number of turns of L1-use from lessons 1-4 and a trend for a higher number, and higher percentage of correct use of the phrasal verb ‘AGREE+PREP’. At lesson 5 (VC₃ original) the trends for LREs and L1-use reverse, and the percentage of correct use of AGREE+PREP drops off. This is the point when the MA intervention took place, changing the conditions of the procedure, so this lesson will be discussed in detail in the following chapter. The study of procedural repetition will be limited to the first four lessons (VC₁ and VC₂) i.e. Lesson 1 vs. Lesson 3.

7.2.1.1 The effect of procedural repetition on language related episodes (LREs)

Kim (2013) found the number of grammatical LREs (LREg) and lexical LREs (LREv) increased in the procedural repetition condition. Kim (2013, p. 12) included pronunciation (LREp) and spelling LREs in her category of lexical LREs. However, there were only two spelling LREs in the current dataset, so they were not considered in this study. My co-rater (Chapter 4) and I disagreed with Kim about including pronunciation in the lexical LRE category. Poor pronunciation may lead to misinterpretation of meaning by the hearer. If this misinterpretation led to an LRE, it was coded as an LREv. However, sometimes students simply asked each other how to say a word or if their pronunciation was correct. In these instances, there was no misinterpretation of meaning, so they were coded as LREp’s. There were only 31 LREp’s in the entire dataset.

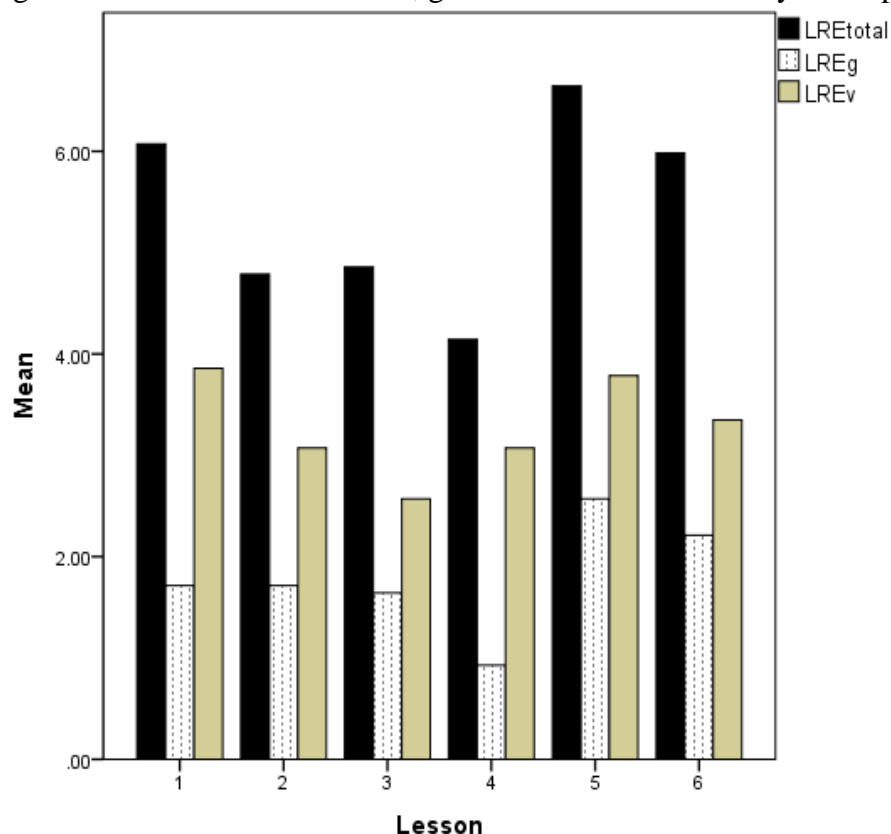
Table 7.2. The total number of LREs, L1-use turns and use of the targeted phrasal verb, “AGREE+PREP” across all teams

Lesson	Task	Language Related Episodes (LREs)				Turns of L1-use			Target phrasal verb: Agree+prep		
		LREg	LREv	LREp	Total LREs	L1-on	L1-off	L1-total	Correct	Incorrect	Total (% correct)
1	VC ₁ original	24.7	54.4	7.0	86.1	88.6	10.2	98.8	5.2	8.8	14 (37.1)
2	VC ₁ repeated	22.9	41.9	0.9	65.7	75.4	50.8	126.2	11.7	6.6	18.3 (63.9)
3	VC ₂ original	22.4	36.8	8.4	67.9	73.8	43.9	117.7	10.7	3.9	14.6 (73.3)
4	VC ₂ repeated	12.9	41.6	3.8	58.3	72.7	60.8	133.5	16.0	6.3	22.3 (71.7)
5	VC ₃ original	35.8	52.7	5.3	93.6	25.9	41.7	67.7	17.2	11.5	28.7 (59.9)
6	VC ₃ repeated	31.0	46.9	6.0	83.8	36.1	19.6	55.7	7.5	5.1	12.6 (59.5)
Original TOTAL		82.9	143.9	20.7	247.5	188.3	95.8	284.2	33.1	24.2	57.3 (57.8)
Repeated TOTAL		66.8	130.4	10.7	207.8	184.2	131.2	315.4	35.2	18.0	53.2 (66.2)
GRAND TOTAL		149.7	274.3	31.4	455.3	372.5	227	599.6	68.3	42.2	110.5 (61.2)

* The shaded regions represent the tasks done as part of the metacognitive awareness (MA) raising intervention.

The data in Table 7.2, and Figure 7.1, show trends in the opposite direction to Kim's (2013) with respect to LREs. Total LREs, LREg's and LREv's all showed an overall decrease from Lesson 1 to Lesson 4.

Figure 7.1. Mean number of total, grammatical and vocabulary LREs per 10min



To confirm the trends in the table and figure, a repeated measures ANOVA was performed on total LREs, LREg and LREv for VC₁ and VC₂. The results are shown in Table 7.3.

Table 7.3. RM-ANOVA results for LREs.

Variable	Correction	df	F	sig	partial eta squared	observed power
LRE total	Greenhouse-Geisser	2.028	2.242	0.125	0.147	0.418
LREg	Greenhouse-Geisser	2.274	1.723	0.193	0.117	0.353
LREv	Greenhouse-Geisser	2.248	1.502	0.239	0.104	0.310

Table 7.3 shows there were no statistically significant findings of the RM-ANOVAs, however, effect sizes were large for LRE-total, and moderate for both LREg and LREv. In other words, there was a large decrease in LREs from Lesson 1 to Lesson 4, although we

cannot claim that it was not a random event. The low observed power of all three tests indicates that a larger sample may have found significant results.

Table 7.2 shows a drop in LREv from Lesson 1 to Lesson 3, from 54 to 37, a procedural repetition difference, which was found to be not quite statistically significant, with a small-to-moderate effect size, $t=2.154$, $df=13$, $p=0.051$, Cohen's $d=0.46$. This, however, is in the opposite direction predicted by Kim (2013) who found lexical LREs rose with both types of repetition.

Pronunciation LREs (LREp) were not normally distributed, so a Related-samples Friedman's 2-way ANOVA was performed. It found a statistically significant difference existed among the LREp across the four lessons, $\text{Chi sq.}=9.387$, $df=3$, $p=0.025$. Table 7.2, however, reveals the differences were between original and repeated versions of the same task, so the results of post hoc comparisons will be reported with the results of task repetition, below.

7.2.1.2 The effect of procedural repetition on L1-use

The use of L1 was not investigated by Kim (2013), however, it was a major concern for the participating teachers in Carless' research (Carless, 2004, 2007, 2009). Kim (2013) included translation as a type of lexical LRE, and in this study, any L1 that was part of an LRE was counted as LRE data, not L1-use data. On-topic L1 was defined as talk in Korean in which the participants talk about the topic in general, elaborate on earlier ideas, comment about their progress, discuss task procedures or organize their activity. Off-topic L1 was defined as talk in Korean about topics unrelated to the VC task. It was predicted that L1-on would decrease over multiple repetitions as participants' English developed, and as their familiarity with the task procedures eliminated the need for organizing talk in L1. Off-topic L1 was predicted to remain relatively stable.

Table 7.2 shows that over the first four lessons, the number of turns of L1 use steadily increased, driven by large increases in off-topic L1, counter to expectations. On-topic L1 had a slight decrease, as predicted. Related-samples Friedman's 2-way ANOVA were performed to investigate whether or not the increase in L1-off and decrease in L1-on were significant. The increase in off-topic L1 was not found to be significant ($N=14$, $\text{Chi sq.}=1.606$, $df=3$, $p=0.658$). Likewise the decrease in L1-on topic talk was found to not be significant ($N=14$, $\text{Chi sq.}=0.541$, $df=3$, $p=0.910$). So, despite the trend in the data for

increasing use of off-topic L1 and decreasing use of on-topic L1, no statistically significant result was found for procedural repetition.

Figure 7.2. The number of turns of L1 use across lessons

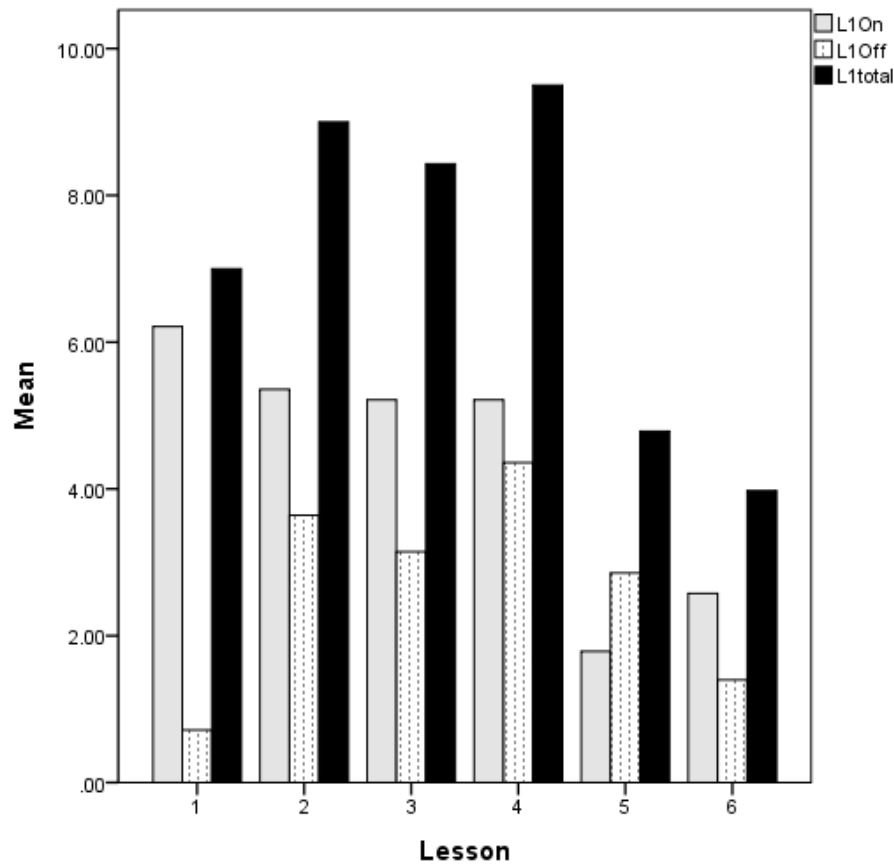


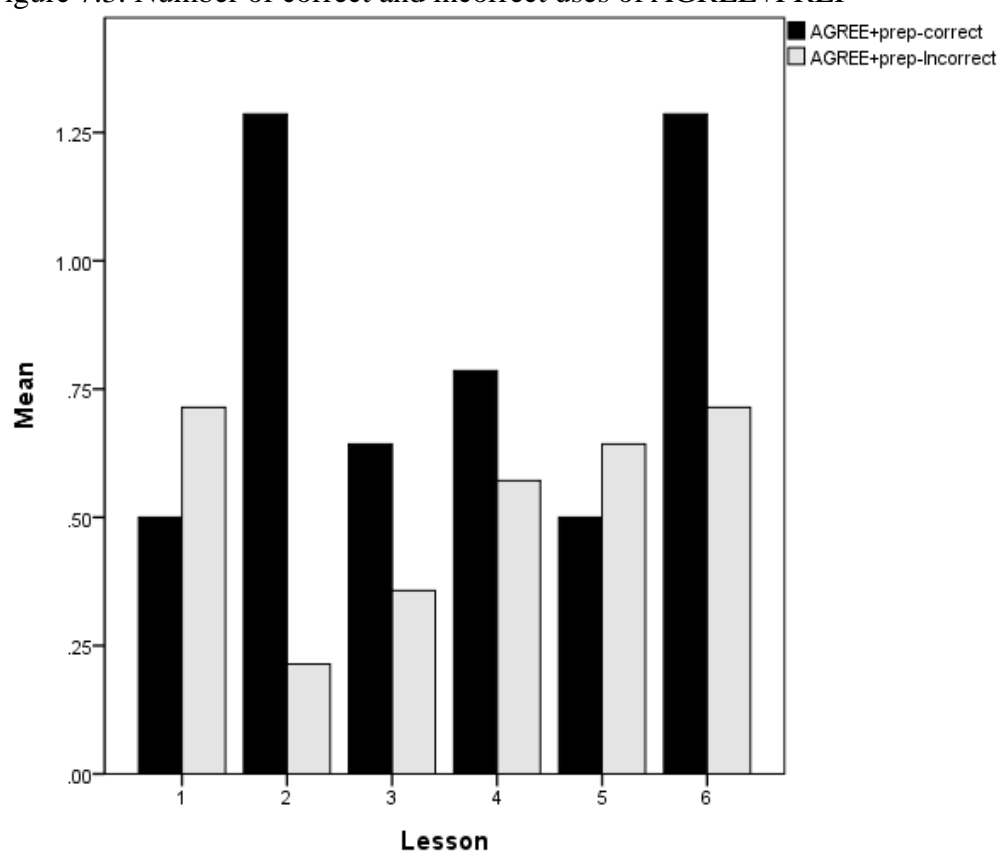
Figure 7.2 shows the behaviour of L1 use across lessons clearly: the upward trend overall from Lesson 1-4, the slight decrease in on-topic L1 across the same four lessons and the upward trend of off-topic L1. The sharp change at Lesson 5 supports the notion that the MA intervention changed the task procedure in ways that task repetition did not.

Tentatively, it is proposed that the rise in off-topic L1 across lessons could indicate general fatigue with the VC task or increasing interpersonal familiarity among the group members. I will return to this point in the qualitative analysis later in the chapter as both pertain to the effectiveness of procedural repetition as a pedagogical intervention (Chapter 8).

7.2.1.3 The effect of procedural repetition on ‘AGREE+PREP’ use

Table 7.2 indicates an overall increase in correct use of the phrasal verb, while incorrect use remained similar across Lessons 1-4. Figure 7.3 confirms this. However, when Related-samples Friedman’s 2-way ANOVA were performed, no statistically significant difference was found ($N=14$, $\chi^2=3.069$, $df=3$, $p=0.381$). However, lessons 1 and 3 show less correct uses than lessons 2 and 4, so there may be an effect of task repetition, which we will investigate in the next section.

Figure 7.3. Number of correct and incorrect uses of AGREE+PREP



7.2.1.4 Summary of procedural repetition

The investigation of procedural repetition was carried out on the first four lessons of the study. I only reported the first four lessons (VC_1 and VC_2) because the fifth and sixth lessons (VC_3) also involved the metacognitive awareness (MA) task and was therefore not an equivalent condition. This is a limitation of this study. There was a planned counterbalance to the MA design (VC_4), which would have presented two separate repetition studies, each of two original and two repeated tasks. However, classroom

constraints, such as national holidays and university-staged events prevented this from taking place.

No statistically significant results were found for procedural repetition. There were some trends in the data, with respect to LREs and L1 use, and these trends were in a direction which suggested negative effects for task repetition. The number of LREs decreased and the number of turns of L1 increased, but these were not statistically significant effects. However, the moderate to large effect sizes imply more than a trend, but the small sample size and low power of the tests did not allow for a statistically significant finding. Both the LRE and L1 data went against predictions. The LREs behaved counter to published reports (Kim, 2013). The L1 data ran counter to predictions due to increased off-topic L1. On-topic L1 reduced as expected. I tentatively propose, and will illustrate with qualitative data in Section 7.3, that both the LRE and L1 off-topic data can be explained by participants' fatigue with the VC task and increasing familiarity among the members of the teams.

7.2.2 The effect of task repetition

It was noted during the investigation of procedural repetition that there may be some differences across 'original' vs. 'repeated' tasks, and it is to these I now turn. It was predicted that if students did the same task twice, the first time they did the task would serve as a model or template which would allow resources to be freed up for the learners to pay attention to form (Bygate & Samuda, 2005). This should lead to more LREs, particularly about grammar (LREg), and there should be less need for L1 use.

7.2.2.1 The effect of task repetition on LREs

To investigate the impact of task repetition each VC task was coupled in a pair of original and repeated versions. It was predicted that the number of grammatical LREs would increase in the repeated task because participants would have more resources to focus on their language. It was also predicted that there would be fewer LREv in the repeated task because learners would remember the meanings they were trying to express. Kim (2013) confirmed these expectations. (The LRE data, reproduced from table 7.2, is in table 7.4.)

Different from Kim's (2013) finding, the total LREs dropped in each case of task repetition. In direct contrast to her findings, the number of LREg, decreased in each original-repeated pair. There was a small decrease from Lessons 1 to 2, a much larger

decrease from Lessons 3 to 4, and another small increase from Lessons 5 to 6. For LREv, the results are closer to Kim's. There was a decrease between Lessons 1 and 2, an *increase* between Lesson 3 and 4 and then a decrease from Lesson 5 to 6. Pronunciation LREs also decreased in two of the three VC tasks, and had a very small increase at VC₃.

Table 7.4. LREs across task repetitions

Lesson	Task	Repetition Condition	Language Related Episodes (LREs)			
			LREg	LREv	LREp	Total LREs
1	VC ₁	original	24.7	54.4	7.0	86.1
2		repeated	22.9	41.9	0.9	65.7
3	VC ₂	original	22.4	36.8	8.4	67.9
4		repeated	12.9*	41.6	3.8	58.3
5	VC ₃	original	35.8	52.7	5.3	93.6
6		repeated	31.0	46.9	6.0	83.8

*statistically significant decrease from the original task

A one-way ANOVA was run to determine if the repeated performances were different from the original performances for LREg's, LREv's and total LREs. For LREp's, a related Wilcoxon Signed Ranks test was done. No significant effects were found for any of the LRE types.

When *post hoc* comparisons were done, there was only a statistically significant decrease from the original to the repeated tasks between Lessons 3 and 4 for LREg (Paired sample t-test, $t=2.687$, $df=13$, $p=0.019$, Cohen's $d=0.56$). The effect size is also moderate. This is in direct contradiction to Kim (2013) who found that task repetition increased the number of LREg and decreased the number of LREv. Overall, there was no statistically significant impact of task repetition on LREs, although there is a tendency that may be confirmed with a larger sample size.

7.2.2.2 The effect of task repetition on L1 use

Repeating the task was predicted to be the cognitively easiest condition for the participants. Because of this, they have more mental resources available to address language form. In addition, since L1-use was reported by the participants to be used for unfamiliar words and concepts (Chapter 6), it was anticipated that learners would use less L1 because they already were familiar with the topic and associated vocabulary. However, this was not the case. Table 7.5 has reproduced the data from table 7.2 with respect to L1 turns.

Table 7.5. Total turns of L1-use per 10min VC task talk

Lesson	Task	Repetition Condition	Turns of L1-use		
			L1-on	L1-off	L1-total
1	VC ₁	original	88.6	10.2	98.8
2		repeated	75.4	50.8	126.2
		CHANGE	(-13.2)	40.6	27.4
3	VC ₂	original	73.8	43.9	117.7
4		repeated	72.7	60.8	133.5
		CHANGE	(-1.1)	16.9	15.8
5	VC ₃	original	25.9	41.7	67.7
6		repeated	36.1	19.6	55.7
		CHANGE	10.2	(-22.1)	(-12.0)

Tasks VC₁ and VC₂ show slight decreases in L1-on (on-topic Korean) and increased L1-off (off-topic Korean) from the original to the repeated version. There is also an increase in total L1 use from the original to the repeated version. Task VC₃ (Lessons 5-6) is opposite. It shows an increase in L1-on and a decrease in L1-off; and L1 is markedly lower than in VC₁ and VC₂. Lesson 5 involved the metacognitive awareness (MA) task so this switch in pattern from original to repeated, and the overall drop in L1 could be the result of the MA intervention (Chapter 8).

However, when the data in Table 7.5 were examined statistically, no significant differences were found for any of the original vs. repeated pairs. Off-topic L1 use showed greater changes in counts of turns per 10 min of talk than did on-topic L1 with large increases at Lesson 2 and Lesson 4, followed by a large decrease at Lesson 6. Table 7.6 gives the results of Related-Samples Wilcoxon Signed Rank Tests for off-topic L1 use.

Table 7.6. Results of Related-Samples Wilcoxon Signed Ranks tests for effects of task repetition on off-topic L1 use

VC task	Variable	Original performance	Repeated performance	W	Standard Error	Asymp. Sig.
VC ₁	L1-off topic	10.2	50.8	47.2	14.3	0.889
VC ₂	L1-off topic	43.9	60.8	29.5	8.4	0.405
VC ₃	L1-off topic	41.7	19.6	13.0	9.8	0.139

On-topic L1 was considered a resource available to learners to help them complete the task. As such, its use was not proscribed. It was predicted that participants would have less need for L1 in the repeated task, as it was their second attempt at the task-topic combination and would be able to remember items they had needed in the original task

performance. On-topic L1 use showed minor decreases in the repeated task of VC₁ and VC₂, but the decreases were not statistically significant. In VC₃-repeated task (Lesson 6), there was a large increase, not statistically significant, compared to the original task (Lesson 5). Table 7.7 has the results of the statistical analysis of on-topic L1 for the three tasks.

Table 7.7. Results of Related-Samples Wilcoxon Signed Ranks tests for effects of task repetition on on-topic L1 use

VC task	Variable	Original performance	Repeated performance	W	Standard Error	Asymp. Sig.
VC ₁	L1-on topic	88.6	75.4	47.5	14.3	0.889
VC ₂	L1-on topic	73.8	72.7	36.0	12.6	0.812
VC ₃	L1-on topic	25.9	36.1	64.0	14.3	0.196

7.2.2.3 Task repetition and ‘AGREE+PREP’

The use of the phrasal verb ‘*agree with*’ was predicted to improve over procedural repetitions, but did not. With task repetition, it was also predicted to improve. As with the other variables, it did not show a clear trend. The raw data per 10 minutes of values clarification task talk is in Table 7.8.

Table 7.8 shows very little in the way of trends in the correct or incorrect formation of AGREE+PREP from original to repeated task. The percentage of correctly formed phrasal verbs increased from original to repeated tasks in VC₁ yet remained equal in VC₂ and VC₃.

Table 7.8. The number of correct, incorrect and total uses of ‘AGREE+PREP’(per 10 min)

Lesson	Task	Repetition Condition	Target phrasal verb: Agree+prep			
			Correct	Incorrect	Total	% correct
1	VC ₁	original	5.2	8.8	14.0	37.1
2		repeated	11.7	6.6	18.3	63.9
		CHANGE	6.5	(-2.2)	4.3	26.8
3	VC ₂	original	10.7	3.9	14.6	73.3
4		repeated	16.0	6.3	22.3	71.7
		CHANGE	5.3	2.4	7.7	(-1.6)
5	VC ₃	original	17.2	11.5	28.7	59.9
6		repeated	7.5	5.1	12.6	59.5
		CHANGE	(-9.7)	(-6.4)	(-16.1)	(-0.4)

In VC₃, the number of correct forms, number of incorrect forms, and the total combined instances all decreased in the repeated task, yet the overall percentage was only 0.4% less.

However, the original task in VC₃ contained the most instances of both correct and incorrectly formed AGREE+PREP, and the repeated version dropped to the least total instances of any of the six Lessons. There was just one statistically significant difference among all the original-repeated pairs: in task VC₃, the incorrect use of AGREE+PREP decreased significantly in the task repetition condition (Related-Samples Wilcoxon Signed Ranks $W=1.00$, $SE=5.9$, $p=0.028$). This may be a function of the MA task, and will be discussed further in the next chapter. The data as it stands suggests that task repetition had no impact on the correct formation and use of AGREE+PREP.

7.2.3 Summary of the effect of repetition on task performance

As with the study of procedural repetition, there were trends but few statistically significant findings. The contrast of the results with respect to LREs when compared with published findings (Kim, 2013) requires explanation, as does the trend of L1 use that countered expectations – more L1 in the repeated condition, particularly more off topic.

The findings of procedural and task repetition hint at worse performance across repetitions in stark contrast with other studies on task repetition (Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005; White, 2011). We now turn to qualitative data in order to present, discuss and evaluate explanations as to why repetition did not have the predicted impact in this case.

7.3 A qualitative perspective on procedural and task repetition

The results presented in the previous section revealed very limited results for task repetition, either procedural or task repetition. Moreover, these results differed somewhat from other research and need explanation. This section addresses this discrepancy through qualitative analyses of task performance talk, interviews, participant talk during the metacognitive awareness (MA) task performance and video evidence. Of particular interest was the manner in which task performances differed across repetitions and how that influenced the quantitative results. For detailed descriptions of classroom talk, two teams were randomly chosen using the random number generator at www.random.org (Haahr, 1998), teams A3 and B3. Portions of their transcripts will serve to illustrate

temporal changes. First, however, I start with interview and MA data, to determine if participants understood the purpose of repeating the task.

7.3.1 Student perspectives on procedural and task repetition

To understand why the results of the repetition interventions did not show the predicted results, I turn first to the voices of the participants and later to their in-task activity. For the most part, their comments about repetition are in line with the instructor's and should therefore have produced results similar to predictions. However, their task performances tell a somewhat different story.

7.3.1.1 Student perspectives on procedural repetition

When interviewed, every participant reported that procedural repetition was beneficial for two main reasons. The first reason was that the values clarification (VC) task developed knowledge of the content, and second was the participants' belief that 'practice makes perfect'.

Procedural repetition aided content knowledge consolidation

The participants' perceptions of the relationship between the values clarification task and content knowledge were illustrated above in Section 5.4.2. When discussing reasons for doing the same type of task many times (i.e. procedural repetition), the participants usually mentioned that the VC task's structure of understand, express, explain, negotiate and rewrite forced them to think about the topic and to present their ideas more logically.

For example, Jon, in his interview, said that he liked to do the VC task many times because he thought it helped him speak more "*specifically*." By this, he meant that doing the VC task developed his ability to concentrate on the "*theme*" (i.e. the topic) of the lessons and helped him clarify his thoughts in English.

Extract 7.1. Jon claims repeating the values clarification develops topical/content knowledge

We just do the strongly agree or disagree this one [the VC task]. Maybe I think it is more good than last semester because we can concentrate the some part of theme so we can know about it and then we can talk about it. And then after CLASS or after this semester when we have a topic for this, something like mass media, we can talk or speak or spe- specially ... yeah specifically. So I think it's good

Jon's perspective was valuable because he was one of the few participants in both phases of the study. He had experienced the first phase in which multiple tasks were used (Chapter 3) as well as the second in which the VC task was used regularly at the expense of many of the others. Jon's report that the VC task was helpful for enhancing his knowledge of the course content and developing his ability to discuss it indicates that he felt a match between the task procedure, the task topic and the textbook content, and he felt that the VC task enhanced his knowledge of the content and his ability to express himself. For these reasons, he preferred a repeated use of the VC task to the use of many tasks.

Procedural repetition as 'practice-makes-perfect'

Other participants claimed that doing the same task repeatedly helped improve their skill of discussion and negotiation. Jessie's saw the repetitions as being a form of 'practice-makes-perfect,' and envisaged the ideal, 'perfect', answer as one similar to the input.

Extract 7.2. Repeating the task is 'practice makes perfect'

Researcher

Um why do you think I did the same job [VC task] so many times?

Jessie

Because I think Dan and Gunther and you [the input script] ... is ... one of the answer paper. So I think ah we have to discussion discuss like that. So, uh we want to make that discussion structure like that so again again we try and try

7.3.1.2 Student perspectives on task repetition

No participant reported negatively about repeating the task. Most said it was a good idea because it provided them the opportunity to think more about the topic and get their opinions more correct. No participant reported that they felt the need to make their opinion or ideas more lexico-grammatically accurate. That is, they did not report language-learning benefits with respect to accuracy.

Task repetition as a chance to fix and improve

Almost every participant, in the interviews and MA task said that repeating the task was

an opportunity to make their response better. Jon (Extract 7.3) said that doing the VC task a second time was helpful. He claimed that doing it again allowed him and his group members the chance to be more creative, that is, to add more interesting content. He did not mention improving the accuracy or grammar of the repeated version.

Extract 7.3. Jon – task repetition allows for creative enhancement

Yeah maybe uh I think for example, first day you give us the script and then we- we do the script and then have a answer this is just a basic sentences. and then you give- give the second one the next day and then ah! we always say 'ah we did it' and then we did something like a basic one and then today we can change more for creative or so

For Jon, the first (original) task performance was a 'basic' version that could be built on. Lucky (Extract 7.4) also felt that the repeated task was a chance to add more and was a chance to fix mistakes. It is unclear from the first sentence of her answer if she meant content of language mistakes, but her follow-up sentence indicates that she meant more content.

Extract 7.4. Lucky – task repetition is a chance to fix mistakes and add more

Researcher

We do it [the VC task] on Wednesday and I have you do it again on Monday. Why do you think I have you do it twice?

Lucky

mm. because when the topic is given for the first time then we could make a lot of mistakes. But when the topic is given again, then we could REMEMBER what we had just said in the last class and we could add more

These two extracts reveal that the students were able to identify reasons for repeating a task similar to the reasons I included repetition in the study – improved performance in repeated versions of the task (Bygate, 2001; Lynch & Maclean, 2000, 2001; Pinter, 2005, 2007). This implies that Lucky and Jon are both in stated alignment with the teacher's workplan (Atkinson et al., 2007). They both point to the potential for lexical and content improvements when repeating the VC task. However, this potential was not fulfilled and ecological, contextual factors need to be examined to determine why (Section 7.4, below).

Task repetition as a missed chance to fix and improve

Despite recognising the potential for improving performance in repeated tasks, students did not engage in the task in a way that would increase learning opportunities. When asked about this, Nada (Extract 7.5) reported she felt that repeating the task provided the opportunity to improve, but said she and their classmates chose not to use that opportunity.

Extract 7.5. Nada and the missed opportunity to improve in the repeated lesson

It's just JUST a little bit different from the last week (lesson) I think from two days or five days between the class we can maybe think and maybe rearrange ... our opinion... we don't really try to but we can.

When asked why she did not try to “think and rearrange” she answered, “*I’m too lazy and it’s too tough.*” In other words, she was not willing to make the effort to improve her performance when given a chance to do so. This would also affect the task performance of her teammates and reduce the learning opportunities for all of them.

Therefore, despite knowing the purpose of task repetition, to improve their task performance, some students simply did not use the repeated task as a chance to improve their performance. In the interviews, they expressed a focus on fixing mistakes or adding more in the repeated task, which is what was predicted. However, the equivocal quantitative results, above, suggest Nada’s answer might have been closer to most teams’ behaviour.

This demonstrates the importance of doing ecologically valid research. Volunteers for after-class experiments will, in general, not be “too lazy” because they were not too lazy to volunteer, thereby creating a sampling bias in such studies. The classroom-based studies on repetition also lacked complete ecological validity in that they: (1) involved repetition in a single lessons (Lynch & Maclean, 2000, 2001); (2) were ‘special lessons’ with a ‘new’ technique that was not an element of regular class procedure (e.g. Kim, 2013); or, (3) happened many weeks after the original lesson (Bygate, 2001). None of the studies reported by the authors were characterised as typical classroom activity done on a regular basis. This study, in which student-volunteers’ participation was, in essence, to show up to class, do their normal class behaviour, and let me transcribe their talk allowed for even the laziest students to volunteer.

Having all (but one) student participate in this study allowed the investigation of a full range of commentary and behaviour. We can see from Nada's comment why this is important. From her self-described, 'lazy student' words, we learn one reason students in ongoing classes do not perform tasks as expected – sometimes they are lazy and take a more pragmatic approach to task performance. Without Nada's participation in the research, something most 'lazy' students may well never do. We would not find this in research that is more traditional.

Reflecting the notion of being too lazy or unwilling to improve task performance the second time around, it was noted that, in some cases, there was very limited interaction in the repeated condition. To investigate, I turn to the task performance transcripts.

7.3.2 Student behaviour in task performance

When transcripts of task performance were reanalysed to determine factors that might have prompted the unpredicted results of task repetition, it was found that over repetitions some teams became less engaged with the task (Platt & Brooks, 2002). No team refused to do the task or engaged in completely off-topic behaviour, but some teams avoided discussion by claiming to be finished or by simply speaking less about the topic. We look at each of these in turn.

7.3.2.1 Lack of engagement: not starting

To be engaged with a task, one has to start doing it. With the values clarification task, if everyone in a team strongly agrees with all the prompts there is no need to discuss or revise anything. As the lessons progressed, some teams tried to avoid the task completely by claiming all members strongly agreed with all prompts. Team B3 is an example of one team that did this. Some excerpts of Team B3's talk from the start of the tasks in different lessons are presented below in Extracts 7.6 to 7.9. The first two extracts come from early in the study and show the group quickly engaging with the VC task.

Extract 7.6. Team B3, Lesson 1(VC₁-original)

Turn	Speaker	Utterance
4	HELEN	number one (reading) separating boys and girls into different schools is too //old fashioned for modern society\\
5	SARA	//old fashioned for modern society\\
6	HELEN	does yours agree ? strongly agree ?
7	SARA	L1 <i>a moleugesseo</i> <i>아 모르겠어</i> (tr=oh I don't know)
8	T	(gives the answer sheet to write answer) you need that for later (goes to another team)

9	HELEN	okay . disagree ?
10	SARA	disagree
11	HELEN	yeah uh I agree I think that it's too old fashioned
12	SARA	I- I think it is NOT old fashioned but
13	HELEN	uh huh
14	SARA	it doesn't work well
15	HELEN	ah: so
16	SARA	so
17	HELEN	maybe we have just take TOO because too is like
18	SARA	too
19	HELEN	is like uh it's VER:Y old fashioned but so we have to take that and it's separating boys and girls into different schools is //old fashioned\\ for modern society

In Lesson 1, (Extract 7.6) the prompt was read aloud. Sara said she does not know, so Helen continued, saying her opinion and then starting to give her reason (turn 11). Sara's turn 12 indicated slight disagreement with Helen (finished in turn 14), but she did not explain fully. At turn 17, Helen attempted to revise the statement. In short, they have gotten on task very quickly. Similarly, 7.7, shows a similarly quick start.

Extract 7.7. Team B3, Lesson 3 (VC₂- original)

Turn	Spkr	Utterance
5	HELEN	uh what do you think about number one ?
6	SARA	(3.0) mm (4.0) what is eliminate
7	HELEN	mm ?
8	SARA	eli: mi: nate
9	SUE	L1 <i>eobs-aeda</i> 없애다 (tr=eliminate)
10	SARA	(repeats) L1 <i>eobs-aeda</i> 없애다 (tr=eliminate) °is not a good idea ° hmm strongly uh disagree ?
11	HELEN	yeah me too why ?
12	SUE	but here it says ° (<i>unint</i>) ° (she is reading the prompt)
13	HELEN	I think that because we are EQUALS and about gender bias words something that uh before long time ago (2.0) eh: (5.0) what do you think ? (to Sue)

Extracts 7.6 and 7.7 show this team is trying to engage the task from the beginning of their allotted time. They used different approaches to starting: reading the prompt aloud (Extract 7.6) and directly asking for the opinion (Extract 7.7) but either approach begins the discussion. They then asked about words in the prompt or started giving their opinions. In the case of Extract 7.6, they were already trying to revise the prompt by the end of the extract – Helen suggested taking out the word 'too' (turn 17).

This contrasts with the next two extracts from later in the study. Extracts 7.8 and 7.9 are

from Lesson 4 and 5 (VC₂-repeated and VC₃-original, no MA condition). Here, they required intervention from the instructor (T) to get started.

Extract 7.8. Group B3, Lesson 4 (VC₂-repeated)

Turn	Spkr	Utterance
1	T	okay get talking guys h
2	SUE	I totally agree
3	HELEN	uh-huh
4	SUE	(12.0) I strongly agree in all
5	T	(17.0) (comes by this group) the easy way is to say hey Helen what do you think
6	SARA	uh Helen what do you think ?
7	T	hmm good question
8	HELEN	I'm agree I strongly agree with ALL of them
9	SUE	(to T) what if we ALL of us agree- //strongly agree in all\\
10	HELEN	//strongly agree in all\\
11	T	you strongly agree with all of them
12	SUE	yeah
13	T	you're going to have to tell me why
14	HELEN	okay why we agree with them ?
15	SUE	becau:se we have to choose . even-
16	T	-you thinking that changing the language will change the culture or do you think changing the culture will change the language
17	SUE	changing the culture
18	T	so then you have to disagree with that
19	HELEN	first we have to change the mind
20	SUE	then eliminating gender biased language will help to change .
21	HELEN	it's gender biased culture or mind maybe the stereotypes that the people have of men of wo-womens and mens (8.0) (to T) this is uh- it this is going to be the style of the: group discussion right ?
22	T	mm probably not
23	HELEN	ah no ?
24	T	I haven't decided
25	SUE	eliminating gender bias will (voice trails off) culture in a positive way

Helen and Sue claimed they agreed with all the statements, and there were two long pauses at the outset. The instructor (T) walked by and prompted them to speak at turn 5. Then when they told him that they agreed with all the prompts, he reminded them of the public report at turn 13 “*you're going to have to tell me why*”. Helen tried again to start their discussion (turn 14) but Sue's reason for agreeing was reduced to “*because we have to choose . even*”, meaning that the only reason she agreed strongly was because they had to choose something, not because she really agreed. The instructor asked them a specific question and when Sue answered, he pointed out that she should have disagreed (turn 18), indicating she had not understood the prompt. However, this was a repeated task, with prompts only slightly different from those used in the previous lesson (Extract 7.7),

meaning that the ‘misunderstanding’ was not really a misunderstanding. Then Helen interrupted and asked about their upcoming test (turn 21), moving their talk off-task again. Twenty-five turns into their discussion and they had not yet expressed an opinion to each other. This difficulty getting started on their task was uncharacteristic of their earlier lessons. Essentially, they had tried to avoid the task by claiming they had, in essence, finished and had no need to talk. In reality, they did not agree with all the prompts. Moreover, Sara had completely misunderstood the meaning of one of the prompts and would never have realised this had she not been pushed to discuss her opinion. The next lesson, a case of procedural repetition, showed similar problems getting started as well (Extract 7.9).

Extract 7.9. Group B3, difficulty getting started, Lesson 5

Turn	Speaker	Utterance
2	HELEN	hh (sighs) what do you think number one ?- about number one
3	KSA	I agree
4	HELEN	agree
5	SUE	hmm
6	T	okay can you-
7	SUE	why don't we change it to strongly agree
8	T	what do you have to change (in the prompt) to make it strongly agree?
9	HELEN	no we STRONGLY agree
10	T	no don't lie (laughter) that's not what's on your paper
11	SUE	mass media is: (6.0) why don't we strongly agree ?
12	HELEN	it's not uh ah now I strongly agree !
13	SUE	mass media is not as bad but .
14	HELEN	it's bad
15	SUE	but NOT helpful

In Extract 7.9, Sue's suggestion at turn 7, spoken directly in front of the instructor, “Why don't we change it to strongly agree?” was spoken as a suggestion to change their response, not as a request for information. T intervened and asked for information, but Helen followed Sue's approach, to change their answer sheets to avoid discussion. She restated that she now strongly agreed again at turn 12, but Sue had already re-engaged with the task and started to discuss changes. Repeating the task had reduced engagement for this team.

The changes to the starts of the task illustrated in Extracts 7.6 – 7.9 demonstrate a change in task engagement. They have tried to avoid doing the full task by strongly agreeing with the prompts thereby removing any need to discuss changing them. Such an approach

would reduce talk in general, LREs in particular, because changes were no longer being discussed, and the time spent not doing the VC task would be an opportunity to talk off topic in their shared L1. The evolution in approach from ‘do the task’ to ‘avoid the task’ in this team and others, may therefore explain the unpredicted quantitative results with respect to task and procedural repetition.

What caused this shift? One possible explanation may well be boredom with the task or the topic or both. Kim (2013) reported that her participating teacher had expressed fears the students would be bored doing the same task three times (in the task repetition condition). However, this study’s combination of task and procedural repetition meant that by the time Team B3 performed Extract 7.9, they had done the VC task five times on three topics. It is possible that boredom extended to procedural repetition as well.

When asked in the stimulated recall interview, if she had been trying to get out of doing the task in Lesson 4, Helen laughed and smiled, then said, “*We didn’t really know why why we were doing like that one so much. I know this uh now ah but then it was like uh (waves hands) again?*” But, in the same interview, she also had expressed preference for doing the same task many times, and not doing many different tasks, indicating at the time of Lesson 5, before the debriefing session, it was not apparent to her why the VC task was used so much. This point will be raised again in the next chapter. However, Helen’s pattern of task engagement across tasks shows that using the same task multiple times does not mean the participants will put the same, or even similar, effort into doing it each time. Therefore, instructors will need to communicate their reasons for repeating tasks and tell students how to benefit from doing them twice (See Chapter 8).

7.3.2.2 Lack of engagement: speaking less

Extracts 7.10, and 7.11, below, are transcripts of Team A3, doing an original task response to a prompt from Lesson 3, and a repeated task response to the same prompt from Lesson 4. The prompt, rather cynically, read, “*If boys and girls go to different schools, the government will probably give more money to the boys’ school.*” Each table is followed by figures containing the written, final version, of their talk.

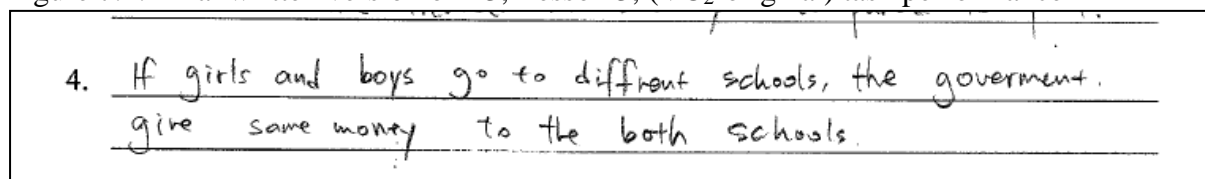
Extract 7.10. Team A3, Lesson 3, (VC₂-original) task performance of the prompt

Turn	Speaker	Utterance
169	SUNNY	//let’s move on four\\
170	WINNIE	(to NANCY) //you lie\\ (laughter)

171	SUNNY	I: s- strongly disagree
172	WINNIE	oh me too me too
173	SUNNY	(to NANCY) you disagree
174	NANCY	same money
175	SUNNY	yeah the government should
176	WINNIE	yes
177	NANCY	between
178	WINNIE	ah the way
179	SUNNY	SHOULD give more money uh- uh
180	NANCY	no-
181	SUNNY	//-give SAME money\\
182	NANCY	// give same money\\
There are 12 turns of off-topic L1 here		
193	SUNNY	... if girls and boys go to different schools the government mm should //give\\ ₁ same money
194	WINNIE	//will\\ ₁ . . uh SHOUL-DUH (=should) give same money //to the\\ ₂
195	SUNNY	//to the \\ ₂ both schools
196	NANCY	the government will probably give more give- uh
197	SUNNY	//money\\
198	NANCY	//same money \\ (1.0) per population
199	SUNNY	too difficult (rejects Nancy's idea outright)
200	WINNIE	same money
201	SUNNY	uh the government give give . same ? money ? to the both school:s .

This talk in the original task started with a suggestion to move on to prompt four followed by all three members' expression of disagreement with the prompt. They then start collaborating on changes by nominating words to include in the final version. Nancy suggested '*same money*', omitting '*the*'. Sunny suggested '*should*', which Winnie agreed to. The talk is a series of nomination and building with non-agreed items simply ignored (turns 177-8). There is collaboration and overlapping speech, with members finishing each other's talk (e.g. turns 179-180). They then move off topic in L1 for 12 turns of talk, about a paper that Sunny is reading for another class. Sunny returned to the topic at turn 193, by stating a revision of the sentence she thought was suitable. The next few turns are characterised by collaboration and simultaneous speech, mostly nominating different modal verb to include: *will*, *should* or *will probably*. They do not develop the idea of different amounts of money based on population distribution (Nancy, turn 198) as this approach was deemed '*too difficult*' by Sunny (turn 199). This had the effect of ending further discussion on the matter and leading to the writing of the final version (Figure 7.4). Note, in Figure 7.4, none of the suggested modal verbs appears in the final version, even though Sunny had said '*should*' and she was the writer, and Winnie had strongly emphasised it at turn 194.

Figure 7.4. Final written version of A3, Lesson 3, (VC₂-original) task performance



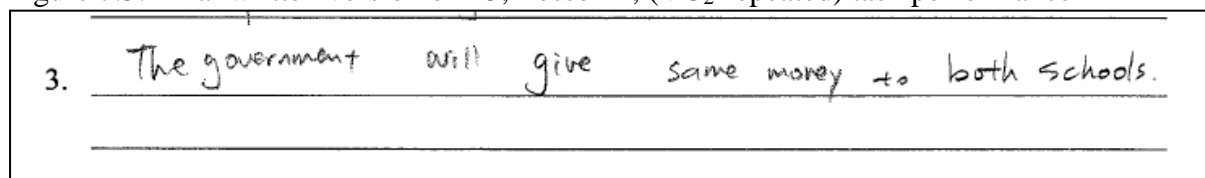
Extract 7.11 contains the transcript of the same team working on the same prompt at the start of the next class. The long pause at the beginning is due to one of the members writing the previous prompt's answer.

Extract 7.11. Team A3, Lesson 4, (VC₂-repeated) task performance of the prompt

Turn	Speaker	Utterance
84	WINNIE	(15.0) next
85	NANCY	I:: strongly disagree
86	WINNIE	//disagree //
87	SUNNY	//me too// disagree
88	NANCY	give same MONEY
89	SUNNY	uh yea //same money//
90	WINNIE	//same money//
91	NANCY	to both schools

The entire process of expressing agreement/disagreement, nominating an answer and agreeing on it took seven turns among three speakers. There was no discussion of the prompt. There was also no attempt at developing the notion of '*money per student*' that Nancy had mentioned in the previous lesson. There was only a brief suggestion of the new statement and they moved on. The revised statement is half the length of that in the original task, and it is missing the if-clause completely. The modal verb, '*will*', has been mined (Boston, 2008) from the prompt. Figure 7.5 has their final written version.

Figure 7.5. Final written version of A3, Lesson 4, (VC₂-repeated) task performance



Extracts 7.10 and 7.11 demonstrate one team's evolving engagement with the repeated tasks. They have a 'finish it quickly' approach. Both Extract 7.10 and 7.11 show a quick move away from expressing their opinions and reasons to creating a revised version of their sentence, similar to team D1(Section 5.2.1). Extract 7.11 shows an impoverished interaction with no negotiating or development of the topic or the new sentence at all.

Unlike Jon’s claim (Section 7.3.2.1), Team A3 did not add to or develop their original task’s ideas; the repeated task was the ‘*basic one*’. We shall see in Chapter 8 that the metacognitive awareness (MA) intervention greatly improved this team’s approach.

7.3.2.3 Task repetition as a ‘second chance to get it right’

Not all teams reduced their amount of talk. Some groups used the repeated task as a chance to improve what they said (Section 7.3.1). Improvement was not limited to content, however, as LREs were sometimes re-enacted in the repeated task because one or more members still were not sure of the item’s meaning or form. An example of this comes from Team B3 – Helen, Sue and Sara. Two extracts from VC₂ are presented in Extracts 7.12 and 7.13, below. The first extract is an extended LREv about the word ‘stereotype’. The LRE involved giving examples, gestures and finally the use of L1 to resolve. The second extract, from the repeated task, shows that their solution was not fully understood by Sara, and they needed to re-do the LRE to clarify.

Extract 7.12. B3, Lesson 3 (VC₂-original), resolving the LREv- ‘stereotype’ first attempt

Turn	Speaker	Utterance
43	SUE	number two what’s stereotype
44	HELEN	when you think- ah for example when you say COOK who cooks ?
45	SUE	the woman
46	HELEN	ha ha ha ha ha that is the stereotype . you think ?
47	SUE	L1 편견? <i>pyeongyeon</i> (tr=prejudice, bias)
48	HELEN	I don’t know it L1 편견? <i>pyeongyeon</i> (tr=prejudice, bias, biased view)
49	SARA	L1 편견 <i>pyeongyeon</i> (tr= prejudice, bias, biased view) cliché

Extract 7.12 exemplifies this team’s approach to the VC task. Like many teams, they started work on a prompt by asking about vocabulary. Here the targeted word was ‘stereotype’. Helen, in turn 43 gives an example – by suggesting that when Sue thinks of the word ‘cook’ it is the woman who cooks. Both of them laugh at this, but the example is correct. Sue suggests an L1 translation, but Helen did not know if the translation were correct. Sara repeated the translation (turn 49) and added another synonym, ‘*cliché*’. Helen then confirmed the meaning with her cell phone’s dictionary.

The next lesson evolved differently to the preceding lesson. They took much longer

getting started (Section 7.3.2.1, Extract 7.8). To recap, the participants had told the instructor they strongly agreed with all the prompts. The instructor reminded them they would still need to say why they agreed in their public report. This forced them to restart their task. Extract 7.13 happened later in the same task, once their talk had gotten underway. They again encountered the word ‘*stereotype*’, and again it caused difficulty.

Extract 7.13. Team B3, Lesson 4, (VC₂-repeated), resolving LREv- ‘stereotype’ second attempt

Turn	Speaker	Utterance
63	HELEN	uh-huh (encouraging Sara)
64	SARA	often times discrimination have no reason . // <i>LI inga</i> ? <i>인가?</i> (tr=is) \\\
65	SUE	//but it says\\ stereotypes
66	SARA	stereotypes <i>LI igeos-eun</i> <i>이것은</i> (tr=this) stereotypes about women . is uh is not from our uh // <i>LI inga</i> ? <i>인가?</i> (tr=is) stereotypes \\\
67	HELEN	//it is\\ uh-huh
68	SARA	I think stereotype MEANS uh um woman- women has to do something ? and women has
69	HELEN	uh-huh
70	SARA	children and women are expected to
71	HELEN	mm-hmm do something
72	SARA	but discrimination some- sometimes discrimination- discrimination have NO
73	HELEN	reason
74	SARA	reason . for example uh uh uh <i>LIgeu-geos</i> <i>그것</i> (tr=that) for example before people- people gave more money for boys school why why why::: why have why have to: ? do-do that ? h* uh the answer is boy is better than- boys matter than b- girls b- this is NOT a stereotype
75	HELEN	yeah it’s a stereotype uh-huh
76	SUE	that’s a stereotype
77	HELEN	that’s a stereotype that the people THINK the yea- gir- //boys are better than girls\\
78	SUE	//boys are better than girls\\
79	SARA	this is also stereotype ? I think stereotype is uh include only behaviour ? no ?
80	HELEN	ah:
81	SUE	I don’t know
82	HELEN	no no no no
83	SUE	I’m not sure hh
84	HELEN	no stereotype is everything
85	SARA	everything
86	HELEN	uh-huh everything
87	SARA	uh-huh ? (<i>unint</i>)
88	SUE	(<i>unint</i>)
89	SARA	ah:
90	HELEN	(<i>unint</i>)
91	SARA	okay then I:: strongly agree number two .

Sara’s talk illustrates one of the stated reasons for repeating the task, to develop a deeper

understanding of the language used in the task. Extract 7.13 shows Sara negotiating the meaning of the word *stereotype*, which had been the subject of an L1-mediated LREv in the original task (Extract 7.12).

Here, in the repeated task, the word *stereotype*, which had been misunderstood by Sue last lesson, caused problems for Sara too. The L1-mediated resolution from Lesson 3, had not clarified what was meant by the word *stereotype*. It took from turns 66 to 85 for Sue and Helen to help Sara understand that a stereotype involves thinking and not only action. This extract demonstrates that resolved LREs are not necessarily learned or completely understood by all members of the team and that repeating the task allowed Sara the opportunity to clarify her misunderstanding.

In addition, the repeated version, Extract 7.13, showed more interactive collaboration (Foster & Ohta, 2005). Helen used continuers at turns 63, 67 and 69 to encourage Sara to speak; there was co-construction of utterances at turns 70-1 and again at 72-3; overlapping speech at turns 64-5, 66-7 and 77-8; and a long effort at negotiating the meaning of *stereotype* from turns 74-84. The entire LRE required 28 turns whereas only eight turns were used in the original.

Sara also used the opportunity afforded by the repeated task to express her opinion and receive feedback on her misunderstanding. Helen and Sue used the repeated task as a means of clarifying their own definition of *stereotype* through helping Sara. Sue, for example, at turns 81 and 83 expressed doubts as to what a *stereotype* was, and it was only Helen's insistence that a *stereotype* could be thinking as well as doing (turns 82, 84,86) that solidified her knowledge of the word. This implies that even if LRE counts, L1-use, and uptake of specified forms did not differ across repetitions, repeated task performance provided a different range of language learning opportunities for different participants.

The examples from this context revealed the extent to which the two task performances (original and repeated) differed, despite having the exact same participants, and that these differences were motivated by the students and not the task. The analysis in this section showed that each task performance has its own life, separate from, yet influenced by, previous performances. This influence is not a direct cause-and-effect influence, as students chose to perform the task in a way that suited their needs at that time. The teams in this section took a pragmatic, finish-quickly approach to the repeated task, not one consistent with using the task performance to improve language use. This suggests that if

we teacher/researchers expect task performance to match the task workplan, and thereby drive language learning, we need to show students what is expected of them and why. The metacognitive awareness (MA) activity reported in Chapter 8 is one way of doing this.

7.4 Discussion of the task repetition results

It was expected that repeating the task would lead to more LREs (Kim, 2013) and there would be more LREs in the procedural repetition conditions than in the task repetition conditions. That is, it was expected that the original tasks would have less LREs than the repeated tasks and there should be a general increase in LREs from Lesson 1-6. It was also anticipated that there would be more L1 use in original tasks and L1 use would decrease from Lesson 1-6 as students became more familiar with the task procedure and their English improved. In addition, it was expected that task-natural grammar, AGREE+PREP would be used with increasing correctness from Lessons 1-6 due to the increasing number of incidental exposures (cf. Webb et al., 2013).

However, each original task (Lesson 1, 3 and 5) all saw higher levels of LREs and less L1-off than their corresponding Repeated tasks (Lesson 2, 4, and 6). The new topic for each original task may have prompted more LREs and less off-topic L1, but it was predicted that there would be an increase in LREs and a decrease in L1-off from Lesson 1-2 (task repetition effect) and Lesson 1-3 (procedural repetition effect) but the opposite happened. The correct use of AGREE+PREP did not change systematically. The difference between the predictions and the results needs to be explained.

In the qualitative results, I suggested boredom as one reason some participants demonstrated unwillingness to get started or went off-topic in L1, particularly in the task repetition condition. Boredom was also reported by the participating teacher in Kim (2013) when she stated a preference for procedural repetition to prevent the learners from becoming “tired of doing the same thing” when using the same content (p.17). Kim found overall superior performance, as expressed by the number of grammatical and lexical LREs, in the procedural repetition condition when compared to the more ‘boring’ task repetition. Task repetition still produced more grammatical LREs in her study, but lexical LREs (similar to LREv in this study) decreased. The present study, however,

found an overall decline in performance measures; decreasing LREs, both LREg and LREv, and increasing L1-off topic, from Lessons 1 to 4 and then a reversal at Lesson 5 to a level slightly superior, to Lesson 1, when the metacognitive awareness (MA) task was done. Lesson 6 saw a decline in LREs from Lesson 5's level. So, Kim's students, even in the more potentially 'boring' condition, task repetition, still used more LREs with increasing repetitions, whereas, the only time in the current study that LREs increased (and L1-off decreased) co-occurred with the MA intervention.

The qualitative examination of Teams A3 and B3 showed that participants did not openly express boredom or unwillingness when not participating, yet their behaviour showed a lack of engagement with the task, and was directed toward a quick and easy completion of the task and not with language learning from the task. Team B3 attempted to avoid doing the task by claiming to be finished before they had started. Team A3 resolved the prompts with very limited interaction and waited until the instructor was not near before talking about off-topic things in L1 (cf. Carless, 2004). Kim (2013) did not examine L1 talk at all, so there is no possible comparison with her study, but the different behaviour in terms of LREs requires explanation.

7.4.1 Constant team membership or changing partners

In this study, team membership was kept constant because I wanted to control for interlocutor variables (Plough & Gass, 1993; Storch, 2002) and maintain a within-subjects dimension. This way, any effect of task repetition could be attributed to repetition and not to the new partner. Therefore, the values clarification (VC) tasks were done with the same members each time, and the participants changed interlocutors when doing other activities during the programme. This gave control over interlocutor variables in the VC task, and gave students the chance to interact with a variety of partners. In contrast, other researchers changed interlocutors with each repetition (Bygate, 2001; Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005, 2007; White, 2011).

I expected that keeping team members constant would alleviate feelings of anxiety and provide learners a safe platform from which to push their language development. However, there was also the possibility that having new interlocutors could make the task 'feel' new, and push learners to be more precise because they were with someone new. The interviewed participants held one of these two, mutually exclusive points of view on the matter and no participant was impartial.

From a ‘familiar-is-safe’ perspective, Jon and Nada expressed strong preferences to stay in the same team because they felt they could express themselves better with familiar members. In her interview, Nada said, *“I really love meeting or hanging out with new people but when it comes to the discussion I want to hang out with familiar friends”* (Nada, interview). To her, being with familiar people meant she knew how they thought and she could feel comfortable expressing herself. Jon claimed that having familiar faces let him talk more: *“Familiar group members encourage me to speak more”* (Jon, interview). From the ‘new-is-a-push’ perspective, Janine and Jessie both urged me to change the partners in the VC groups next time.

Extract 7.14. Janine wanted a variety of interlocutors

I wanted that we change the group members [for the VC task]. I wanted to know more about what the another people’s thinking and talking more”.

Janine strongly disliked keeping members the same because she was curious about their ideas, and wanted to talk to others. Jessie, however, raised a point pertinent to the issue of the unpredicted results.

Extract 7.15. Jessie – the same group members increases L1 (Korean) use

I think uh this curriculum is good but actually we need to change member many time ... So, uh I mean, ... Yes I always talk about some topics always with maybe Cathy and Lucky and maybe that make uh fun situation and using Korean situation so, it uh it interfere that [learning]

Jessie’s claim that familiarity with her team members promoted the use of Korean raises a possible explanation for the rise in off-topic L1 across repetitions. It also implies that many of the benefits of task repetition found in the literature (e.g. Lynch & Maclean, 2000, 2001; Pinter, 2005; White, 2011, see Chapter 2) could be due not so much to the act of repeating *per se*, but to repeating to someone new. When in a group of familiar people, learners may be more likely to engage in off-topic talk because they know their interlocutor well, and because they know each other well, they have more things to talk about. With a new partner, this would not be the case. This means that, in the present study, keeping group members together over the study made them more familiar with

each other, and this familiarity may have enabled them to go off topic. This explains the off-topic L1 increases observed in some teams, but says nothing about LREs.

The different LRE counts in my study compared to Kim (2013) can be accounted for by considering what learners remember when learning in groups. They remember not only the language or content but also information about their interlocutors' personality, their opinions, and their language proficiency. Because learners remember their partners' opinions they may feel bored more easily, as they are not doing something new. This would lead to off-topic L1 and less LREs. Even those who stay on task will use this knowledge they have about their partners during their repeated task performance to avoid language and topic-related difficulty, not to create more (cf. The Cooperative Principle, Grice, 2000).

Therefore, keeping team membership constant reduced the linguistic and conceptual challenge of the task repetition conditions. The changes in the language of the prompts between original and repeated tasks were small (Section 4.4.1) and learners could rely to a certain extent on their memory of the previous interaction. Because they were familiar with their partners, participants may have had an idea, in advance, of what their interlocutors' opinions were likely to be, thus making the repeated task an opinion-gap with less 'gap' than the original.

The participants would also know what language their group members already knew; knowledge that would become *increasingly accurate* across repetitions as learners interacted with each other more and more. There would then be less need for LREs, as the participants would keep the talk in their partners' linguistic 'comfort zones'. Thus a shared memory of the VC task mediated the interaction in the task repetition condition (cf. Bygate & Samuda, 2005). In Kim's (2013) study, because each partner was new each time, the participants could not employ a shared memory of the other person doing the task to facilitate their interaction. They would then need to re-negotiate the level of language in the repeated task performance, driving up the number of LREs. This question should be investigated in the future.

There may be other issues related to having new partners. Ewald (2004) reported that Ned, a student who had been struggling, switched seats in order to change the way groups were formed, thereby giving him new interlocutors. Ewald reported that this benefitted Ned's development of L2 Spanish by providing "a different collaborative environment" (Ewald,

2004, p. 173). Ewald's report on Ned and a higher-level student, Amanda, identified affective reasons behind their desire to change groups, similar to Janine's curiosity about others (Extract 7.14). The present study, through examining the difference in results with those of Kim (2013), suggests that there may also be cognitive benefits to changing the membership of groups and some of these lie in the lack of shared experience with the other interlocutors doing the task.

This has practical implications for teachers working in settings where classroom layout is not conducive to changing partners (e.g. sitting in rows of pairs of students, such as in many South Korean classrooms) and in which the noise of moving from partner to partner might disturb neighbouring classes. In addition, teachers who allow students to self-select partners or to make partners based on who was seated nearby (e.g. Philp et al., 2010) may wish to rethink this policy. The results of this study, when combined with the previous literature suggest that teachers make the effort to move students to new partners when repeating a task or a task procedure.

7.5 Chapter summary

This chapter investigated the effects of procedural and task repetition on task performance. The results were not in line with published research: LREs decreased, L1-use increased and there was no evidence of acquisition of AGREE+PREP. The qualitative analysis showed that participants understood why the task was repeated, but during the task performances, sometimes their in-class activity was not aligned with maximum language learning potential. Participants judged procedural repetition to be useful but after the third repetition (Lesson 3) some teams' performance dropped off. Teams B3 and A3, exemplified the manner in which this happened.

In the stimulated recall interviews, most participants understood the reason for task repetition and the benefit they would get from doing things a second time. However, at the time of the task performance, even participants who knew and understood the reasons for repeating the task and liked to repeat the task did not always make a consistent effort to follow the workplan. For example, Team B3 used an avoidance strategy, claiming to be unable to do the task because they all agreed when they had not fully understood the task

or did not really agree with the prompt. It took the instructor's intervention to keep them working. When they finally got started on the repeated tasks, their task performances were enhanced vis-à-vis the originals. They showed more collaborative interaction and they resolved language problems that persisted from the original task. Team A3 by contrast, spent less time and fewer turns developing their opinions and final revisions to the prompts across the first four lessons. To fill in the unused time, they spoke in L1 on off-topic subjects. This was dramatically changed for Team A3 after the MA intervention, which shall be investigated in the next chapter.

CHAPTER EIGHT: THE EFFECT OF A METACOGNITIVE AWARENESS (MA) ACTIVITY ON TASK PERFORMANCE

8.1 Introduction

This is the fourth and final chapter of the results and discussion. It reports on two lessons, Lessons 5 and 6 (VC₃). The previous chapter presented data and analysis on the impact of task repetition on task performance. It was found that task repetition had an effect opposite to predictions. The qualitative analysis suggested a number of reasons for the unexpected results: (1) participants had become familiar with each other, which encouraged L1 use and inhibited pushing linguistic performance; and, (2) participants used task completion strategies to finish quickly – strategies that enabled them to avoid discussing the prompts fully. The qualitative analysis also found that task repetition helped participants review and even re-learn items previously engaged with in LREs. Data presented in Chapter 7 (e.g. Figures 7.1, 7.2 and 7.3) suggested changes in task performance occurred at Lesson 5 as a result of the metacognitive awareness (MA) intervention reported in this chapter.

This chapter reports on the final intervention of the study, which involved raising learners' metacognitive awareness (MA) of the purposes of doing the task and how to learn through task performance. The MA task engaged students with their own learning process and encouraged them to generate ideas for approaching tasks in order to make the most of the learning opportunities provided. The MA task was expected to improve task performance while at the same time providing a data source to shed light on learner cognition. It was a survey task that required students to analyse their classroom situation and learning behaviour through discussion and comparison with other participants (see Chapter 4). Unlike Ewald (2004) who included skits and other activities based on the instructors' ideas of student task performance, I limited my role to making the survey and providing "my version of the answer" to the students *after* groups had finished the task and subsequent presentations.

As in the two previous results chapters, I present quantitative data first and follow with qualitative data to explain and support the numeric data.

8.1.1 Implementation of the MA task

The MA task was done in Lesson 5 of the study. There were two groups, a treatment group, Group 1 (classes A and D) and a control group, Group 2 (classes B and C). Group 1 was given the MA task at the start of Lesson 5. After reading the questions and thinking about their own answers, they were moved into new teams (not the VC task teams) to discuss their answers. After 30 minutes, they chose a representative to present a summary of their talk to the class. After all groups had presented, the instructor gave what he termed were 'possible' answers to some of the items. All task talk was recorded and later transcribed for analysis. Then the participants moved back to their values clarification (VC) task groups and performed the third VC task (VC₃-original). They did VC₃-repeated task at the beginning of the next lesson.

Group 2 did not do the MA task as part of the study; rather, they did another listening activity from their textbook. They did the MA task later, after the study ended, as part of the debriefing session. This arrangement allowed for a between-subjects comparison for the effects of the MA task along with a within-subjects comparison of task performance using previous lessons as a baseline.

8.1.2 Research questions

The research questions addressed in the chapter are as follows.

1. What is the effect of the metacognitive awareness (MA) activity on VC task performance for groups who had the MA activity when compared to those who did not? (i.e. What are the across-groups effects of the MA activity?)
 - a. Does the MA task change the number of LREs?
 - b. Does the MA task change the amount of L1 used?
 - c. Does the MA task change the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?
2. What is the effect of the MA activity on task performance for groups who had the MA activity when compared to performances in earlier lessons? (i.e. What are the within-groups effects of the MA activity?)
 - a. Does the MA task change the number of LREs?

- b. Does the MA task change the amount of L1 used?
- c. Does the MA task change the correct use of a task-natural phrasal verb (Loschky & Bley-Vroman, 1993)?

Because it was found (see Chapter 6) that no use of targeted items from the input processing activities took place, the input condition was not factored into the analysis.

8.2 Quantitative data and analysis

When examining potential effects of the MA activity, I investigated the effect of the MA activity on the treatment group, Group 1, compared to the control group, Group 2 (between-subjects), and also on the same group before and after the treatment (within-subjects). I start by presenting the results of the between-subjects investigation.

8.2.1 The between-subjects effect of the MA task

The first research questions (1 a, b and c) referred to the differences in VC task performance between groups that received the MA treatment and those that did not. The overall results are reported in Table 8.1, which shows Lesson 5, in which the MA task was performed and the Lesson 6, the repeated version of the task in order to determine if there were delayed benefits from the MA task.

Table 8.1. Between-subjects comparison of the number of LREs, turns of L1 and formation of AGREE+PREP across MA conditions

Variable	Lesson 5 (VC3-original)			Lesson 6 (VC3-repeated)		
	Group 1	Group 2		Group 1	Group 2	
	MA+ condition	MA – condition	Difference ‡	MA+ condition	MA – condition	Difference
LREg	20.8	14.9	5.9	17.1	13.9	3.2
LREv	33.3	19.3	14	26.9	20	6.9
LREp	3.2	2.1	1.1	3.3	2.6	0.7
Total LRE	57.3	36.2	21.1	47.2	36.5	10.7
L1-on	16.1	9.8	6.3	16.2	19.9	-3.7
L1-off	16.1	25.6	-9.5	5.3	14.2	-8.9
L1total	32.2	35.4	-3.2	21.5	34.2	-12.7
A+P correct †	12.4	4.8	7.6	3.9	3.6	0.3
A+P incorrect	2.7	8.9	-6.2	0.8	4.3	-3.5
A+P Total	15.1	13.6	1.5	4.6	8	-3.4
A+P % correct	82.3	34.9	47.4	83.4	45.7	37.7

† A+P = formation of the phrasal verb AGREE+PREP

‡ Difference = MA+ minus MA- , i.e. groups that received MA – those that did not

8.2.1.1 The between-subjects effect of the MA task on LREs

It was predicted that the MA task would encourage participants to engage more fully with the language-learning purpose of the task and thereby prompt more LREs. Table 8.1 shows that in Lesson 5 the groups that did the MA task (MA+) prior to doing the VC task engaged in more LREs in their VC task than those that did not do the MA task (MA-). This was still true in Lesson 6, but the difference reduced, implying the benefit of the MA task was not limited to the lesson in which the MA task was performed. Independent samples t-tests comparing the MA+ group to the MA- group were performed. The results are in Table 8.2.

Table 8.2. Independent samples t-tests comparing LREs across MA condition

Lesson	Test variable MA + vs. MA –	df	t	p	Effect size Cohen's d	Effect strength
Lesson 5 (VC3- original)	LREg	12	1.000	0.337	0.536	medium
	LREv	12	1.555	0.147	0.831	large
	LREp	12	0.397	0.698	0.208	small
	TOTAL LRE	12	1.558	0.145	0.833	large
Lesson 6 (VC3- repeated)	LREg	12	0.664	0.519	0.355	small
	LREv	12	0.824	0.426	0.440	small
	LREp	12	0.263	0.797	0.141	none
	TOTAL LRE	12	0.876	0.398	0.468	small

None of the comparisons was found to be statistically significant. However, the Lesson 5 data show the effect sizes for LREv and Total LRE were large, over 0.8; and LREg had a moderate effect size, over 0.5. This suggests that in the same lesson as the MA task was performed, total number of LREs and LREv's was a large degree higher, and the number of LREg's was a moderate amount higher, in the MA+ group than in MA- group. We cannot eliminate the possibility that the effect was random, due to the high p values, but the effect was in the predicted direction. A larger number of teams would have been preferable. By Lesson 6, the effect sizes had reduced, indicating that any impact the MA task had, was no longer strong but still present.

8.2.1.2 The between-subjects effect of the MA task on L1 use

Table 8.1 shows that in Lesson 5 the groups that did the MA task prior to doing the VC task used less overall L1 in their VC task. They used slightly more on-topic L1, but more importantly for this study, the amount of off-topic L1 was reduced by 37%, from 26 turns in MA- to 16 turns in MA+. This indicates an increased orientation to learning English through the task performance.

Overall (i.e. total) L1 use was still lower in the MA+ condition relative to the MA- condition in Lesson 6. However, *on-topic* L1 had decreased in the MA+ condition, and was now less than in the MA- condition. Moreover, *off-topic* L1 remained 37% less in the MA+ condition. Overall Total L1 use was also 37% less in the MA+ group. This indicates that the benefit of the MA task, with respect to L1 use, extended beyond one lesson.

However, the results of Mann-Whitney U Tests revealed no statistically significant differences across MA+ and MA- conditions for any of the variables. Indeed p-values were very large. The results of the tests are in Table 8.3.

Table 8.3. Results of Mann-Whitney U tests comparing L1 use across MA condition

Lesson	Test variable MA + vs. MA –	U	p
Lesson 5 (VC3- original)	L1 on	24.0	0.946
	L1 off	24.5	1.00
	Total L1	24.0	0.948
Lesson 6 (VC3- repeated)	L1 on	17.5	0.365
	L1 off	20.5	0.596
	Total L1	18.5	0.442

The indication from Table 8.3 is that any impact from the MA intervention was moderate. Still, turns of off-topic L1, in both Lesson 5 and Lesson 6, were less in the MA+ condition, as predicted, so it appears that the MA intervention had some effect toward reducing the amount of off-topic talk in L1.

8.2.1.3 The between-subjects effect of the MA task on AGREE+PREP formation

Table 8.1 shows that in Lesson 5 the groups that did the MA task prior to doing the VC task had more formations that were correct and fewer formations that were incorrect of AGREE+PREP. Both of these benefits were nearly eliminated in Lesson 6, however, indicating a temporary change. The largest positive difference across groups was the correct formation of AGREE+PREP in Lesson 5, where MA+ had twelve and MA- only five. And the largest negative difference was the incorrect formation of AGREE+PREP in Lesson 5, where MA+ had three and MA- had nine (See Table 8.1). This was in the predicted direction. However, in Lesson 6, the difference in correct formations was almost zero and the difference in incorrect formations now only three. Table 8.4 shows the results of the Mann-Whitney U tests comparing the correct, incorrect, and percentage of correct AGREE+PREP formations across MA+ and MA- groups. As with L1, there were no statistically significant differences to report.

Table 8.4. Independent samples t-tests comparing LREs across MA condition

Lesson	Test variable MA + vs. MA –	U	p
Lesson 5 (VC3- original)	A+P correct	19.5	0.495
	A+P incorrect	12.5	0.990
	A+P % correct	5.0	0.711
Lesson 6 (VC3- repeated)	A+P correct	21.0	0.602
	A+P incorrect	16.0	0.173
	A+P % correct	1.5	0.361

Therefore, we can tentatively conclude that the MA task had limited impact on AGREE+PREP formation, and what effect it may have had was short lived, limited to the VC task immediately following the MA task.

8.2.1.4 Summary of between-subjects effect of the MA task

There were no statistically significant effects of the between-subjects comparison of MA+ and MA-. Effects sizes revealed trends in LREs that were in the same direction as predictions; the MA task prompted more LREs in a subsequent task performance. Likewise, off-topic L1 also followed predicted patterns but not to a statistically significant degree. The use of the task-natural lexico-grammatical item ‘*agree with*’ showed only a small impact, from raising metacognitive awareness, which was not apparent in the repeated task.

We turn now to the within-subjects effects of the MA task to determine if the metacognitive awareness (MA) raising task changed the way participants performed the values clarification (VC) task compared with previous performances.

8.2.2 The within-subjects effect of the MA task

The other main research question pertaining to the MA intervention asked whether it improved task performance of participants compared to their earlier performances of the VC task.

Table 8.5 contains counts per 10 min talk of LREs, L1 turns and occurrences of AGREE+PREP for Lessons 4, 5 and 6. The table shows Groups 1 and 2 as MA- in Lesson 4 because Group 1 received the MA intervention at the start of Lesson 5. The MA survey was not repeated in Lesson 6, but rather Lesson 6 served as a test for delayed effects of

the MA intervention.

Table 8.5. Within-subjects comparison of Lesson 4 task performance with MA+ and MA- conditions from Lessons 5 and 6 on all variables

Variable	Lesson 4		Lesson 5		Lesson 6	
	Group 1 (MA-)	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)
Total LRE	32.8	25.5	57.3	36.2	47.2	36.5
LREg	6.6	6.2	20.8	14.9	17.1	13.9
LREv	23.4	18.2	33.3	19.3	26.9	20
LREp	2.8	1.0	3.2	2.1	3.3	2.6
L1total	50.3	83.2	32.2	35.4	21.5	34.2
L1-on	26.5	46.3	16.1	9.8	16.2	19.9
L1-off	23.8	37.0	16.1	25.6	5.3	14.2
A+P correct †	1.2	14.8	12.4	4.8	3.9	3.6
A+P incorrect	0.0	6.2	2.7	8.9	0.8	4.3
A+P Total	1.2	21.0	15.1	13.6	4.6	8
A+P % correct	100	70.2	82.3	34.9	83.4	45.7

† A+P = formation of the phrasal verb AGREE+PREP

Table 8.5 shows the last three lessons of the study: The lesson immediately prior to the MA intervention, and the two immediately after it. This provides a basis for comparison of the pre-post effects of the MA task on values clarification (VC) task performance.

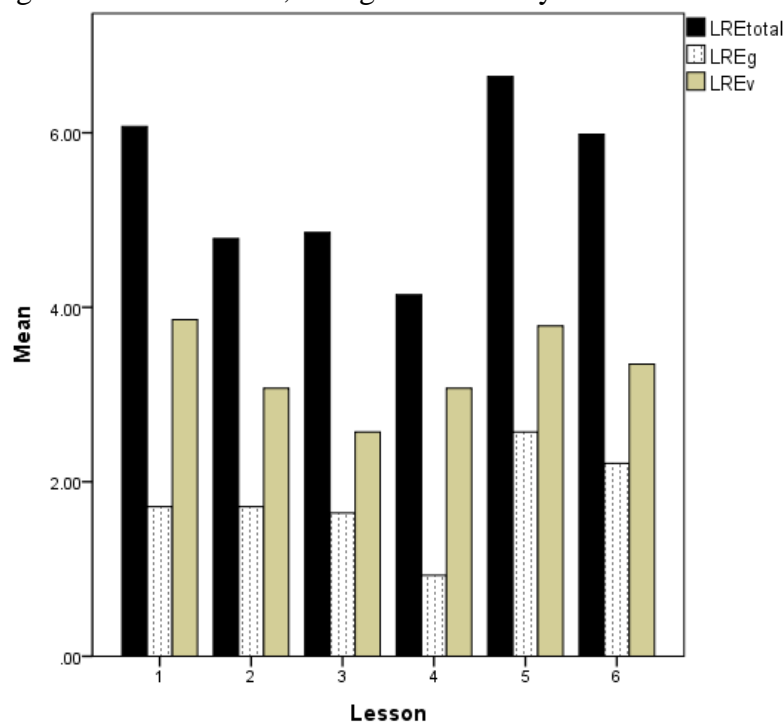
8.2.2.1 The within-subjects effect of the MA intervention on LREs

It was predicted that the MA intervention would increase the number of LREs used in the subsequent values clarification (VC) task because the MA task made learners aware of the language learning purpose of the task, and LREs represent learning in action.

Examining the top row of Table 8.5, the total LREs for Group 1, who received the MA intervention, rose from 33 to 57 between Lessons 4 and 5, and decreased to 47 in the repeated task (Lesson 6). Both Lessons 5 and 6 had more LREs than Lesson 4 as predicted. In addition, the decrease from Lesson 5 to 6 is in line with the task repetition data reported in Chapter 7. However, Group 2's total LREs also increased, from 26 to 36 and remained steady at 36 in the repeated task. In the case of Group 2, LREs rose from Lesson 4 to 5 but not to the extent that they did for Group 1. This indicates a potential effect of changing topic acting in conjunction with the MA intervention. Figure 8.1, a reproduction of Figure 7.1, shows the trends of LREs across all lessons of the study.

There appears to be a slight rise in LREs between Lesson 2 and Lesson 3, another point of topic change. However, the rise in LREs for Group 1 (MA+ group) was about 40% higher suggesting an increase in LREs.

Figure 8.1. Total LREs, LREg and LREv by lesson



To investigate omnibus differences between Group 1 (MA+) and Group 2 (MA-) over the course of the study, i.e. from Lesson 1 to Lesson 6, a RM-ANOVA was used. The within-subjects factor was LRE count and the between groups factor was MA condition. Three RM-ANOVAs were run, one for Total LREs, one for LREg and one for LREv. Table 8.6 summarises the results.

Table 8.6. RM-ANOVA results for LREs: total, grammatical and vocabulary

Measure	Correction	df	F	Sig (p)	Partial Eta Squared	Effect size Category	Observed Power
LRE total	Greenhouse-Geisser	3.211	2.279	0.091	0.160	large	0.549
LREg	Greenhouse-Geisser	2.752	2.698	0.066	0.184	large	0.579
LREv	Greenhouse-Geisser	2.899	1.060	0.377	0.081	moderate	0.258

There were no statistically significant effects across the six treatments despite large effect sizes found for the overall LREs and LREg. These two variables have p-values that lie within Larson-Hall's (2010) claim that $p=0.10$ is sufficient for education research, and given the large effect size in both cases, we can claim that trends in the data show participants focused more on the formal accuracy of their talk by engaging in grammatical LREs after being given the MA+ task.

The RM-ANOVA was conducted over all six lessons. However, the shape of the graph in Figure 8.1 suggests that the first four lessons were trending differently to the last two. If we consider Lesson 4 to be a baseline of current task performance captured immediately prior to the MA intervention, RM-ANOVA including only Lessons 4, 5 and 6, should reveal the temporally localised impact of the MA intervention. LREs were the within-subjects measures and MA condition was the between subjects factor. Table 8.7 shows the results of these tests.

Table 8.7. Results of RM-ANOVA for LREs including only Lessons 4 – 6

Variable	Correction	df	F	Sig (p)	partial eta squared	Effect size Category	observed power
LRE total	Greenhouse-Geisser	1.944	4.286	0.027*	0.263	large	0.681
LREg	Greenhouse-Geisser	1.503	8.667	0.004*	0.419	large	0.888
LREv	Greenhouse-Geisser	1.622	0.506	0.573	0.040	small	0.115

The RM-ANOVA in Table 8.6 included four lessons (Lessons 1-4) in which the between-subjects variable was not a factor because they only had MA- conditions. Eliminating Lessons 1-3 from the analysis, because they did not contribute to the distinction of MA+ vs. MA-, demonstrates more clearly the impact of the MA intervention on task performances immediately before and after it.

The results trend shown in Table 8.7 is in the same direction as Table 8.6: total LREs increased in Lessons 5 and 6 when compared to pre-MA task performances. The increase was due to an increasing number of LREg's and not to changes in LREv's. The effect sizes for LREg's and total LREs were very large.

Post hoc, paired-samples t-tests, comparing Lesson 4 to Lesson 5 were run. The results are in Table 8.8. Total LREs and LREg's were found to be statistically significantly

different, whereas LREv were not. Effect sizes were large and in the predicted direction: LREs increased in the Lesson when MA+ took place.

Table 8.8. Paired samples t-tests comparing Lesson 4 vs. Lesson 5

Test	df	t	Sig (p)	Cohen's d	Effect size
Total LREs	13	-2.740	0.017*	0.83	large
LRE g	13	-3.296	0.06*	1.15	large
LREv	13	-0.856	0.408	0.33	small

* statistically significant result

Table 8.9 shows the results of post hoc comparisons of Lesson 4 with Lesson 6 to determine any longer term effects of MA+ condition, within groups. The results show that there was still a significantly higher number of LREg's in Lesson 6 than Lesson 4, and the large effect size supports this claim. Total LREs were no longer statistically different (at 5% level) but had still increased over Lesson 4, with a medium effect size. LREv's were no longer different.

Table 8.9. Paired samples t-tests comparing Lesson 4 vs. Lesson 6

Test	df	t	Sig (p)	Cohen's d	Effect size
Total LREs	13	-2.052	0.061	0.692	medium
LREg	13	-3.212	0.007*	1.04	large
LREv	13	-0.526	0.608	0.150	none

* statistically significant result

However, between groups contrasts done with the RM-ANOVA verified the finding of Section 8.2.1.1 that there was not a statistically significant difference across MA+ and MA- conditions, although there was a larger increase in all types of LREs in the MA+ condition (Table 8.5). Both Group 1 (MA+) and Group 2 (MA-) increased LREs in Lesson 5, not just the MA+ group, implying that there was an effect of topic (Mass Media) in Lesson 5 that also stimulated LREs. Yet, the within-subjects increase was larger in the MA+ group, so we can claim that within groups task performance improved due in large part to the MA intervention, but with an effect also from topic.

8.2.2.2 The within-subjects effect of the MA task on L1 use

The between-subjects analysis in Section 8.2.1.2 showed that L1 use, particularly off-topic L1, was less in the groups that received the MA+ treatment, although not to a statistically significant degree. It was also predicted that L1 use would decrease within groups after the MA+ treatment.

Table 8.10, reproduced from Table 8.5, shows the turns of L1 use across the Lesson immediately preceding the MA intervention (Lesson 4), the one immediately after the MA intervention (Lesson 5) and the repeated task done at the start of the next lesson (Lesson 6).

The bottom row of Table 8.10 indicates that for the MA+ condition (i.e. Group 1) the total L1 used in the VC tasks decreased from 50 turns prior to the MA intervention, to 32 turns immediately after the MA intervention finished and continued to decrease to 21 in the repeated task, one Lesson after the MA intervention took place. Off-topic L1 dropped more dramatically over the three Lessons. Lesson 4, prior to the MA intervention, had 24 turns, Lesson 5, immediately after the MA intervention, had 16 turns. Lesson 6, the VC₃-repeated lesson, had only 5 turns of off-topic L1. This starkly contrasts to the effects of procedural and task repetition found in Chapter 7, in which off-topic L1 typically increased in the repeated lesson. Although these results were not found to be statistically significant this large decrease of off-topic L1 in the lesson following the MA intervention, is encouraging.

Table 8.10 also shows that the MA- group also had a decrease from 83 turns in Lesson 4, to 35 in Lesson 5, without the MA intervention. The number of turns levelled off at 34 in Lesson 6. The decrease continued into Lesson 6 for the MA+ group, but did not for MA-.

Table 8.10. Number of turns of L1 talk for Lessons 4, 5 and 6 per MA condition

Variable	Lesson 4		Lesson 5		Lesson 6	
	Group 1 (MA-) †	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)
L1-on	26.5	46.3	16.1	9.8	16.2	19.9
L1-off	23.8	37.0	16.1	25.6	5.3	14.2
L1 total	50.3	83.2	32.2	35.4	21.5	34.2

† Group 1 is classified as MA- in Lesson 4, because the MA+ intervention did not take place until Lesson 5.

To determine if there were any statistically significant results, Related-Samples Friedman's 2-way ANOVA were performed for Total L1, on-topic L1 and off-topic L1. None was found to be statistically significant.

With respect to total L1 use, the MA+ condition's trend is in the predicted direction, a decrease in L1 use, but MA- also decreased between Lessons 4 and 5, but not after Lesson 5. One explanation for this would be that the MA intervention did not have an immediate effect, but a delayed one. There may also have been a topic-related effect,

because Lesson 5 was a new topic. Figure 7.2, above, indicated a very slight reduction in L1 between Lesson 2 and 3, when a new topic was introduced. The new topic might have contributed partially to both MA+ and MA- groups' reduction of L1 in Lesson 5, but the further reduction in Lesson 6 could be attributed to the MA intervention.

In addition, one team in the MA- condition used substantially more L1 in Lesson 4 than it did on average throughout the other five lessons. Team C2 was an anomaly, in that they used an average of just over four turns of L1 per 10 minutes throughout the study, except in Lesson 4, where they suddenly used 26 turns of L1 (per 10 min). In Lessons 5 and 6, Team C2 used much closer to their average, five and six total L1 turns respectively. This created a random increase for Group 2 (MA-) at Lesson 4, which inflated the overall total to 83 turns. This random increase then exaggerated the decrease in Lesson 5. No other teams were far outside their regular L1 use at Lesson 4. If Team C2 had used their average amount of L1 in Lesson 4, the decrease in L1 at Lesson 5 would have been similar to that of MA+. This demonstrates one of the limitations of the small size of the study – the impact of random behaviour changes.

Another factor was that some teams used very little L1 (see Table 6.8, above) and there was no possibility for them to use less L1, thereby mitigating the impact of MA on L1 use. In other words, there was a floor effect restricting any quantitative results. Effects of the MA intervention on L1 use in individual teams will be examined in the qualitative analysis, below.

8.2.2.3 The within-subjects effect of the MA task on AGREE+PREP formation

The correct formation of the phrasal verb '*agree with*' had been predicted to improve in the MA+ condition. Table 8.11 gives the totals per 10 minutes of VC task talk.

Table 8.11. Number of turns of L1 for Lessons 4, 5 and 6 per MA condition

Variable	Lesson 4		Lesson 5		Lesson 6	
	Group 1 (MA-)	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)	Group 1 (MA+)	Group 2 (MA -)
A+P correct †	1.2	14.8	12.4	4.8	3.9	3.6
A+P incorrect	0.0	6.2	2.7	8.9	0.8	4.3
A+P Total	1.2	21.0	15.1	13.6	4.6	8
A+P % correct	100	70.2	82.3	34.9	83.4	45.7

In the MA + condition, correct formation of AGREE+PREP increased from Lesson 4 to 5

then decreased again in Lesson 6. This was as predicted. However, incorrect formation rose slightly in the MA+ condition from Lesson 4 to 5, and then dropped again. The rise was not predicted. In the MA- condition, correct formation decreased across all three lessons, which is against predictions made in Chapter 7 about procedural repetition. Incorrect formation initially increased and then decreased.

Related-Samples Friedman's 2-Way ANOVAs were run for correct and incorrect formation per MA condition across the three lessons, and no statistically significant results were found. The results are in Table 8.12.

Table 8.12. Related-Samples Friedman's 2-Way ANOVA for correct formation of AGREE+PREP per MA condition

Condition	Variable	Chi Sq.	df	Asymp. Sig (p)
MA+	Correct AGREE+PREP	4.4133	2	0.127
	Incorrect AGREE+PREP	2.00	2	0.368
MA-	Correct AGREE+PREP	0.000	2	1.000
	Incorrect AGREE+PREP	3.364	2	0.186

The raw data, combined with the Friedman's tests imply that changes in AGREE+PREP were not influenced by the MA intervention.

8.2.2.4 Summary of within-subjects effects

The main within-group effect of the MA intervention was on the number of LREs, particularly grammatical LREs. These increased as predicted with the MA+ group. But they also increased, to a lesser extent, in the MA- group, implying an effect of topic change in addition to the MA intervention. L1 use did not show any statistically significant differences within groups. One team in the MA- group was found to have unexpectedly used substantially more L1 than normal in Lesson 4, when no intervention took place, and this may have had an impact on the overall results. However, The MA+ group did trend in the predicted direction: less overall L1 in the subsequent lessons, and a great deal less off-topic L1. The task-natural grammar form 'AGREE+PREP' was not impacted by the MA intervention. We turn now to a qualitative analysis of the MA intervention.

8.3 Case study analysis and discussion of MA intervention

This section provides a qualitative examination of the impact the MA intervention had on values clarification (VC) task performance. It will present an ecologically valid case study of one team, including their task performances from before the MA intervention and after it, and data from their MA talk. This team will then be compared with other teams. This case study illustrates how the team's performance changed in the lessons after the MA task.

It was noted above, that L1 use could not decrease for some groups, as they were using very little L1 already (floor effect). It is also reasonable to consider there may be ceiling effects on LREs, i.e. the participants might not be able to include more LREs into their talk, because there was no time or need. For the case study, I selected one team, Team A3 because they used the most L1 (Table 6.8) and therefore this team would be least likely to suffer floor effects for L1. Team A3 also used many, but not the most LREs, so they would be not vulnerable to floor or ceiling effects for LREs.

8.3.1 Team A3

Team A3 consisted of Sunny, Nancy and Winnie. They were all female first year students at the university, majoring in three different subjects. No member of this team had lived or studied outside of Korea. None of them ever missed a lesson or arrived late to class.

8.3.1.1 Task performance prior to the MA intervention

Prior to the MA intervention, Team A3 was the team who used the most L1, and the most off-topic L1. They also engaged in the second-most LREs. Table 8.13 shows their total LREs and turns of L1 use (per 10 min of talk) over the entire study (all numbers are rounded to the nearest whole number). The MA intervention took place at Lesson 5.

Table 8.13. Team A3 Total LREs, total L1 and L1-off topic use

Variable	VC ₁		VC ₂		VC ₃ (MA+ condition)	
	Lesson 1	Lesson 2	Lesson 3	Lesson 4	Lesson 5	Lesson 6
Total LREs	6	11	7	6	12	14
Total L1	41	8	26	25	0	2
L1-off	2	2	14	17	0	0

The table shows that Team A3 engaged in more LREs after the MA condition than in any lesson prior to it. It also shows the complete elimination of L1 Korean from their task performances in Lesson 5, and only 2 turns (on-topic) in Lesson 6. By comparison, Lessons 1 and 2 had exhibited little off-topic L1, but a considerable degree of on-topic L1. This indicates that the team was capable of doing the VC task without the need to venture off topic, and after the first Lesson, could do most of their work in L2.

However, in Lessons 3 and 4 a great deal of both on- and off-topic L1 and slightly fewer LREs were found. Team A3's performance in Lessons 3 and 4 were presented and analysed above in Section 7.3.2.2. Portions of this team's talk can be found in Extracts 7.10 and 7.11. These extracts showed a decrease in engagement (Platt & Brooks, 2002) i.e. less discussion in the repeated task (Lesson 4) which left the participants more time for off-topic talk. It was argued in Chapter 7 that in Lessons 3 and 4, they took an approach to complete the task quickly, avoiding discussion and elaboration of their ideas from the previous (original) lesson and this took away opportunities to learn. The lower number of LREs and very high amounts of off-topic talk in L1 shown in Table 8.13 are indicative of this approach.

The aim of the MA intervention was to have participants discuss the language learning process and their participation in and learning through tasks. The MA intervention was predicted to change task performance by increasing LREs and decreasing L1, particularly off-topic L1. Table 8.13 confirms this is what happened to Team A3.

8.3.1.2 Talk during the MA intervention

To investigate how and why the MA intervention guided Team A3 toward a learning-orientation to task interaction, the members of the team's opinions recorded during the MA intervention will be used, in conjunction with the voices from members of other teams who did not change in the expected direction.

The MA intervention

The metacognitive awareness (MA) intervention was designed to get students to think about the instructors' purposes for doing the VC task and, about how they were expected to learn during task performance. In the MA intervention, participants were assigned to new groups with different members. They were required to discuss ten survey questions (see Figure 4.2, Section 4.4.3) about the VC tasks and about how they performed them.

One member of each group presented a summary of their talk to the class. After all the groups' had presented the instructor led a plenary question and answer (Q&A) session for the participants to ask him about the task or learning in general. The MA intervention reflected Ewald's (2004) study in that an implicit approach was taken on the assumption that starting a dialogue among and then with the participants about the task would lead them to a greater awareness of the learning opportunities that arise during the task. This dialogue was expected to guide the participants to make changes to their behaviour that would increase the quality and quantity of learning opportunities in the task (cf. Crabbe, 2003, 2007).

Discussion in the MA survey about L1 use

Participants were asked directly about when and why they used Korean (L1) (Question 8 of the MA survey). As they were with new partners, some participants' responses did not match their actual behaviour. It is not known if they censored their answers or if they did not realise how much L1 they really used, but the mismatch was most true of the heavier users of off-topic L1. They reported that they only used L1 a little, and in what might be described as 'appropriate' or 'teacher-approved' ways. As the heaviest users of off-topic L1, our case study, Team A3's members were no different.

In the MA survey, Sunny was in a new team with Lucky and Nada. Lucky, from a low L1-off group, said, "*even though I cannot speak well in English ... I just say [speak] slowly and try to explain ... just in my extent only in my extent ... not fluently,*" indicating that she tried to use English first, before using L1. Lucky then said she used L1 so that "*other Korean students know what I want to talk [say] so they will help me*" (Lucky, MA survey). After this, Sunny said, "*When we don't know about the verb or ... word's meaning we have to say Korean to explain it to other members*" (Sunny, MA survey). The implication was that she used L1 only to fill in for a lack of L2, similar to Lucky's view, even though the recordings of Sunny's task performances had many instances of her talking off-task in L1. Sunny did indeed explain and translate items for the other two members using L1 as indicated by her response in the MA survey, but by omitting her off-topic L1, she has aligned with the more standard norm of classroom behaviour reported by Lucky (Atkinson et al., 2007).

Nancy said she used L1 "*when I want to explain something*" (Nancy, MA survey). This was true, Nancy often used L1 to explain her opinion (counted as L1-on topic) when she

had trouble expressing herself, but like Sunny, she did not mention to her new team that she used L1 a lot to talk about things unrelated to the task.

Winnie, the third member of A3 took a different stance, *“I think I always use Korean because I can't speak English very well so I don't know some expressions in English”* (Winnie, MA survey). Here, she positioned herself as being less competent in L2 and therefore justified in using more L1. Her MA group member, Jessie, also from a low-L1 group, rejected this by complimenting Winnie: *“If I question to you in English you [always] answer to me in English!”* (Jessie, MA survey). Jessie’s praise of Winnie’s English both established the social norm of the group, to use English when possible, and created a supportive atmosphere among them. Jessie repositioned Winnie as an ‘English user’, a position she took back to her team for the following VC task.

On the other hand, Michelle, a member of Team A2, the second heaviest L1-use team, clearly reported she used Korean. She said that when her friends were in her group she used a lot of L1 because *“Korean just comes up, POP POP POP!”* (Michelle, MA survey). Here she expressed the idea that L1 is for use with friends. No person in her MA group challenged this view or presented an alternative in line with the task-as-workplan, e.g. use English first and use Korean only when you cannot use English. Team A2’s L1-off use slightly *increased* after the MA survey.

Winnie, Nancy and Sunny all returned to their team after this activity and, as reported in in Table 8.13, used no L1 off topic. Their transcripts revealed they only used Korean in L1-mediated LREs, and engaged in far more L2 talk than in previous lessons. Even before they started their VC task they talked about using only English (Extract 8.1, below). It appears that aligning with peer-norms that promoted L2 use during the MA survey encouraged L2 use in the following VC task. In Michelle’s case, aligning with a view that L1 was the tool for friends encouraged subsequent L1 use.

After the MA survey, when the three members of Team A3, the heaviest off-topic L1 users, performed the VC task for Lesson 5, there was a profound change in language behaviour. The three students adapted their responses in the MA intervention to meet the social norms of their new group, ones in line with the teacher and task-as-workplan; and they then changed their task performance to match what their peers had said. This a case of language socialisation because they have become integrated into the new academic setting through talk with other members (Duff & Kobayashi, 2010; Duff, 2010).

Other high L1-off users, who did not hear or adapt to peer-created, classroom norms in the MA intervention, did not undergo a change in language behaviour. This happened despite the instructor publicly promoting the ‘L2 first’ approach in the plenary after the MA survey ended. This suggests the mechanism of the MA intervention was one of peer-originated language socialisation and not instructor-originated norms.

8.3.1.3 Values clarification (VC) task performance after the MA intervention

Post MA: Orientation to L1

It was shown in Table 8.13 that Team A3 went from being a high L1-using team to a low L1-using one in Lesson 5, after the MA intervention. There was an overall decrease in L1-use among MA+ groups, and this change continued into Lesson 6 as well, but Team A3 showed an exceptionally large decrease. The process of change began in the MA survey talk and continued as they wait for the VC task to begin. Extract 8.1 shows their talk immediately after they reform into Team A3, having just finished the MA intervention.

Extract 8.1. Team A3, between MA+ and doing VC task in Lesson 5

Turn	Speaker	Utterance
1	WINNIE	hi girls ! (3.0) hi girls !
2	NANCY	hi hello (moving and sitting down) (8.0) those pencil cases are very similar
3	SUNNY	mm ?
4	NANCY	you both pencil cases //are very\\ similar
5	SUNNY	//OH:\\
6	WINNIE	what the ?
7	SUNNY	uh it's pretty
8	WINNIE	thank you
9	SUNNY	I like this one better
10	WINNIE	L1 <i>hwaiteu ga issne</i> 화이트 가 있네 (tr= you've got white correcting tape)
11	NANCY	use ENGLISH !
12	SUNNY	OH
13	WINNIE	you have a white
14	SUNNY	yeah
15	WINNIE	white tape
16	SUNNY	and you have a fresh (freshener)
17	WINNIE	oil paper
18	SUNNY	AH ! oh my god ! (laughter)
19	WINNIE	be we . don't .
20	SUNNY	mm
21	WINNIE	use Korean
22	SUNNY	okay (laughs) absolutely okay
23	WINNIE	let
24	SUNNY	someone use Korean ? she has to
25	WINNIE	what ?
26	SUNNY	buy: some drinks
27	WINNIE	mm no

28	SUNNY	(laughter) okay
29	WINNIE	no
30	SUNNY	I I think so too
31	T	okay read these quietly (handing out the sheets)
32	SUNNY	I'm so tired
33	WINNIE	me too hh
34	NANCY	read what ?
35	WINNIE	(when paper is passed to her) thank you
36	SUNNY	(5.0) ah strongly agree task (10.0) uh let's do it

There is a warm greeting among the members, perhaps due to being kept in one group for this one task. Nancy commented on the others' pencil cases. When Winnie used L1 to talk about '*white*', the Korean term for 'white correcting tape' (turn 10), Nancy told her to use English, emphasising English. Nancy had never done this in previous lessons, rather, when Korean was used, she would typically follow suit (cf. Jessie's report in Extract 6.16). For the moment, the new alignment co-constructed through social processes in the MA survey remained with her. Winnie agreed and later told the others they will not use Korean that class, keeping the positioning as an 'English user' provided by Jessie in their MA talk. They do not extend the orientation to punishing Korean use (turns 24 – 28), that is, they see value in only using English (L2) but have decided not to devalue Korean.

This talk all took place prior to the VC task. They were using English for social talk and for organising their team, which they rarely did previously. Moreover, during their VC task, team A3 only used two words of L1, both in LREv, items that were not included in L1 counts (Extract 8.2). Therefore, after the MA intervention, Team A3 had repositioned L1. It had once been their medium of social talk, but L2 had now replaced it, leaving L1 only in its teacher-and-peer-approved role as a resource for learning L2. This elimination of L1 reflects an orientation to the task as something to learn English from, and not as something to finish.

Extract 8.2. Group A3's L1 use in Lesson 5

Turn	Speaker	Utterance
151	SUNNY	but they: study hard than us they go to academies or private- vate (/vert/) °lecture°
152	NANCY	mm <i>L1 gwaoe</i> ㄱㅅ ? (tr= personal tutoring at home)
153	WINNIE	//they don't have \ any time
196	WINNIE	bu:t I think uh they: young people don't know how to use the information they get is because of the: k- quantity of information not because of they are stupid
197	SUNNY	quantity (/kan ti ti:/) ?
198	NANCY	quality ?

199	SUNNY	quantity quantity
200	WINNIE	L1 <i>yang</i> ㄅ (tr= quantity)
201	NANCY	(2.0) ah:

Post MA: Engagement with the VC task

The decrease of L1 would not be helpful if the participants did not engage in the task at all. Extract 8.3 below shows them working on Prompt 2. The most prominent change was an overall increase in talk. To discuss one prompt in Lesson 3 (Extract 7.11) Team A3 used 39 turns (plus 12 off-topic turns in L1) and used nine turns to discuss the repeated version of the prompt in Lesson 4 (Extract 7.12). Extract 8.3 shows Team A3 discussing one prompt immediately after the MA intervention. The prompt read: “*Young people are becoming more and more stupid because they watch TV too much and don’t know how to use the information they get.*” The extract is 125 turns long. Admittedly, number of turns is a crude measure, but the increase is striking nonetheless; four times the number of turns used to discuss and revise one prompt after the MA task than before it. By way of comparison, this group used 238 turns to do all five prompts in Lesson 1, one prompt here warranted over half that total.

Extract 8.3. Lesson 5, Team A3, working on prompt 2 after MA intervention

Turn	Speaker	Utterance
125	WINNIE	(15.0) and question two I disagree
126	SUNNY	me too
127	NANCY	agree
128	WINNIE	agree why ?
129	NANCY	(4.0) agree !
130	WINNIE	why ?
131	NANCY	why you both disagree ?
132	SUNNY	no ! nowadays young people h* well- know well how to use the information THEY: are good at usin:g hand- telephones or-
133	WINNIE	-cell phones- (LREv)
134	SUNNY	-cell phones //or internet \ and
135	NANCY	//NO ! \ not not not that mean uh information
136	SUNNY	yes
137	NANCY	ge- uh get the information ? and USE
138	WINNIE	mm mm
139	SUNNY	MM!
140	WINNIE	yes
141	SUNNY	it is . a problem to use this information ?
142	NANCY	uh (=yes)
143	SUNNY	oh: but STUpid is the other (voice trails off)
144	WINNIE	(2.0) yes I- I don’t like this stupid

145	SUNNY	mm stupid and
146	NANCY	uh I heard- but actually when we survey the: times survey the times that we read newspapers or that
147	SUNNY	mm
148	NANCY	most of us say that oh we don't like to wa- see newspaper- read newspapers and (LREv)
149	SUNNY	mm ?
150	NANCY	and some are
151	SUNNY	but they: study hard than us they go to academies or private-vate (/vert/) °lecture°
152	NANCY	mm <i>LI gwaoe</i> ㄱㄱ ? (tr= personal tutoring at home) (LREv)
153	WINNIE	//they don't have \\ any time
154	SUNNY	// a LOT of things\\
155	SUNNY	they do a lot of things to: h* study or: using information ? but (2.0) °but°
156	NANCY	I think what we- if you: want to- uh if you have the interest in the real world's then you can make the time
157	SUNNY	they watch TV too much? nowadays ? they nev- they can't watch TV
158	NANCY	(3.0) oh
159	WINNIE	so what do you talk about just newspaper ?
160	SUNNY	reading newspapers ?
161	NANCY	uh huh
162	WINNIE	for what ?
163	NANCY	um ah ah: for example uh the TV proglam- program (LREp)
164	WINNIE	yes yes
165	NANCY	the uh uh funny program
166	WINNIE	uh uh okay okay
167	NANCY	the funny program and
168	WINNIE	entertainment
169	NANCY	news program and the sur- survey or the- this argue program
170	SUNNY	mm
171	NANCY	uh usually argue program or news desks (nightly news?) are the- uh percent are very LOW
172	SUNNY	mm
173	WINNIE	yes
174	NANCY	but uh bu:t uh . uh but but however ? the fun programs or that dramas are: be- very high //percent\\ of watching
175	SUNNY	//mm:\\
176	WINNIE	okay okay
177	SUNNY	but that is not only young people
178	WINNIE	uh what ? what this mean what this sentence mean ?
179	SUNNY	(3.0) °students are becoming more ° (voice trails off)
180	WINNIE	because they watch TV too much and they don't know how to use the information they get .
181	SUNNY	mm
182	NANCY	ah ah these days very- a lot of informations are floating
183	WINNIE	mm yes
184	NANCY	but yes we always look cell phones and hear the rumours
185	WINNIE	mm okay
186	NANCY	very much- we heard a lot of information . uh compared to . before
187	WINNIE	yes

188	NANCY	ancients (?) but we can't use
189	WINNIE	uh:
190	NANCY	but because we have to divide that- the informations are correct or not
191	SUNNY	mm:
192	NANCY	and we don't have- we: don't have the: spot- spot or the- . we don't have the DAY to: uh talk about this information ? hh
193	SUNNY	oh I THINK-
194	WINNIE	-yes you- you are right uh there are too many information so they: they are difficult to:- to //divide \\
195	SUNNY	//bias\\
196	WINNIE	bu:t I think uh they: young people don't know how to use the information they get is because of the: k- quantity of information not because of they are stupid
197	SUNNY	quantity (/kan ti ti:/) ? (LREv)
198	NANCY	quality ?
199	SUNNY	quantity quantity
200	WINNIE	<i>Ll yang</i> ㄅㄣˊ (tr= quantity)
201	NANCY	(2.0) ah:
202	T	(had been listening from a distance, now approaches) quantity yes . amount
203	WINNIE	yes amount (accepts T's word)
204	SUNNY	ah they are lack of the in- quality information ?
205	WINNIE	(2.0) don't know- young people don't know how to use the information they get
206	SUNNY	mm
207	NANCY	uh
208	WINNIE	it is because of am- amount of in-
209	SUNNY	-mm
210	WINNIE	a lot of information //not because of they are stupid\\
211	SUNNY	//there are a lot of information\\
212	NANCY	yes I think so
213	SUNNY	they are so smart
214	NANCY	yes
215	WINNIE	but there are TOO many
216	SUNNY	-and THEY: can't watch tv well
217	WINNIE	yes I think so
218	NANCY	ah so there are a lot of information-
219	WINNIE	yes so they: so they //don't know\\
220	SUNNY	//ah!\\
221	NANCY	just we re-correct this sentence OR this sentence ?
222	WINNIE	yes
223	NANCY	I think young people's very smart these days but they are smart in . //SOMething\\
224	SUNNY	//only\\ subject math or English
225	WINNIE	okay okay SO young people don't know how to use the information they get
226	SUNNY	mm
227	WINNIE	because there are too many information ? okay ? no ? (LREg)
228	SUNNY	too much information
229	WINNIE	too much information ?
230	SUNNY	mm
231	NANCY	ah:
232	WINNIE	°media ° (voice trails off as she writes) (22.0) ah ? okay (30.0) okay and

		number three // agree \\\
--	--	---------------------------

The engagement with the task in Extract 8.3 compared to Extracts 7.10 and 7.11 (Sect. 7.3.2.2) improved dramatically. The participants actively listened and encouraged each other to talk more. For example, Winnie and Sunny used continuers to keep Nancy fully expressing her ideas from turns 182-192. They expressed disagreement with what their interlocutors said and provided reasons. For example, Nancy at turn 156 disagreed with Sunny and Winnie, and then Sunny defended herself at 157. In addition, co-construction was frequent. For example, Sunny says ‘*bias*’ simultaneously to Winnie’s ‘*divide*’ (turns 194-5). This shows Sunny was attending to Winnie and tried to anticipate what she was saying. There are also examples of other correction (i.e. LREs) about vocabulary. Among them were: Turns 132-4 (*cell phone*), Turns 197-203 (*quantity, amount*), and Turns 151-2 (*private tutoring*). Two self-corrections appear: Turn 148 (*read newspapers*) was an LREv searching for the collocation, and Turn 163 (*program*) was a rare LREp (pronunciation). The only LREg happened, as mentioned in Chapter 5, as they were writing (Turns 227-230, *too much information*). Again, talk about grammar was stimulated while they were writing.

In contrast, the extracts from Lessons 3 and 4, showed very little of any of this interactive collaboration. There were points when they finished each other’s words or overlapped, and a running series of changed modal verbs. However, the modal verb was eliminated in the final version and LREs never occurred. The one point in Extract 7.10 when a participant tried to expand the talk, at Turn 198, when Nancy said “*per population*”, the expansion was rejected, Sunny claimed it was “*too difficult*”, denying them all the chance to negotiate what was meant, speak more, and learn. Her focus there was on finishing, here, after the MA intervention, it was on expression and discussion.

Team A3 had been the biggest user of L1, and was one of the teams with the least developed responses to some of the prompts. As shown in Extract 8.3, completing the MA survey, which prompted reflection, led to changes to task performance. The team members, themselves, noticed and commented on their own performance near the end of the task (Extract 8.4).

Extract 8.4. Group A3 spontaneous self-assessment in Lesson 5

Turn	Speaker	Utterance
335	SUNNY	ah ! so tired !
336	NANCY	(5.0) ah today we talk very: uh
337	SUNNY	mm

338	NANCY	we are ar- (to others) you are very quality high quality
339	SUNNY	high quality °I think° hh

Sunny comments that this talk had made her tired. Nancy then started to comment on the talk, but changed her approach to talk about her members, calling them “*high quality*”. This word, ‘quality’ was part of an LREv (lines 196-204, of Extract 8.3, above). This indicates that they are now engaging with the task from a learning perspective by using words they did not have available for use before the task. It is also a clear example of the members of the team using words they had negotiated earlier, for their own needs, in later talk. This suggests that although teacher-generated input was not used (Chapter 6), self and peer-generated input might be more readily available for use. Tailored post-tests, based on the LREs each team negotiated might be one way to shed further light on this matter (cf. Newton, 2013).

8.3.1.4 Summary of Team A3

The case of Team A3 illustrated the potential impact of the MA intervention. It allowed the participants to reflect on their performance and then become socialised by their peers into a way of doing the (VC) task that was more in line with the task’s aim and that would lead to better learning outcomes (cf. Duff & Kobayashi, 2010). The MA intervention did not have such dramatic effects on other groups, partially because many of them were more talkative and used less L1 prior to the MA survey so they had ceiling and floor effects. The quantitative data above were equivocal, but this case study illuminates how less proficient or otherwise underperforming participants can benefit from such metacognitive awareness activities. In the case of this team, the peer-delivered metacognitive reminder of the purpose of the task, how to perform the task and the role of L1 in it proved to be an effective push toward a more engaged approach to the task.

8.4 Chapter summary

This chapter investigated the impact of a metacognitive awareness (MA) raising intervention on subsequent task performance. The MA intervention was similar to that of Ewald (2004) in that it took an implicit approach. This study used an interactive task – a

survey, in which participants talked to each other about the task. Thus, it used the words of the participants in interaction with other participants as both the intervention, and one of the analysis methods. A positive reason for doing it this way was that peer-peer interaction socialised behaviour patterns, particularly L1 use, to a more language-learning orientation. A negative was the MA intervention depended on peers to do this. In the case of Michelle, she received no peer-pressure to change, and did not.

The quantitative analysis looked at the effects of the MA intervention. The MA intervention had the greatest impact on the number of LREs. Although between-subjects differences were not statistically significant, there was a large effect size in the predicted direction, that is, more LREs in the MA+ condition. Within-subjects analyses also showed more LREs in the MA+ lesson than in previous lessons, these were statistically significant and had moderate to large effect sizes. This result suggests that the MA intervention oriented the participants to focus on language during task performance to a greater extent than they would otherwise have been, and that participants now treated task interaction as a means to learn and clarify understanding of L2 English and not as something to be finished.

A point to be noted is that the within-subjects effect was largest with respect to the immediately preceding lesson, the fourth time the participants had done a VC task. In Chapter 7, I suggested that participants had become bored or they had used their implicit understanding of each other's language level to expedite task completion. When compared to the first lesson of the study, the impact of the MA intervention is not that large (see Figures 7.1 and 8.1). Therefore, the MA survey may have a greater effect later in a programme when participants have some familiarity with each other, or are not certain why they are doing tasks, as a re-focusing tool. However, this needs to be verified.

Quantitative analyses also showed that L1 use changed in the predicted direction, both between and within groups, but there were no statistically significant effects. Qualitative analysis demonstrated that random events, for example, Team C2's sudden increase in L1 at Lesson 4, might have played a role. In addition, some teams did not have room to decrease L1 use, meaning floor effects were also a factor in the quantitative data.

The correct and incorrect formation of AGREE+PREP was not impacted by the MA intervention.

A case study of Team A3, the team that used the most L1 and engaged in many LREs prior to the MA intervention was provided. The analysis demonstrated a dramatic reduction in L1, to zero turns during the VC task. This was argued to be the result of peer-peer talk in the MA survey itself that oriented the three members to using L1 for solving a lack of L2 knowledge and not for other things. Members of teams that were heavy L1 users who did not receive the same message from peers did not show the same change. Team A3 also showed a transformation in the amount and type of engagement during the VC task after the MA treatment.

Despite the lack of statistically significant results on some measures, the trends were in the predicted direction, the benefits of the MA intervention were clearly illustrated in the case of Team A3.

A final point made in this chapter was that these results were found because the learners were in their situation of learning, i.e. an ecologically valid setting. The MA survey resembled their VC tasks in many respects – read, evaluate and express an opinion. This was normal classroom activity, just on a new topic. Students could engage with the task easily and honestly. It is very unlikely to get students to admit their true behaviour to the teacher in an interview, but with their peers, they opened up. As we saw above, the reaction from their peers to what they said affected subsequent VC task behaviour. For example, Jessie said Winnie used English well, and subsequently used much more English at her team; or, no one challenged (no reaction) Michelle's statement about inappropriate use of Korean and her L1 use rose slightly. Relying on the impact of peer comments about task performance to improve task performance requires participants to be real peers – that is, to have a relationship that has emerged over time, which can serve as the basis on which students can give and receive feedback. Such information could only come from an ecological study.

CHAPTER NINE: CONCLUSIONS

9.1 Overview

This study was an investigation of the influence of task implementation conditions on task performance in Korean EFL classrooms. It was carried out in order to address concerns with the use of tasks in EFL settings, particularly: the limited use of L2, student overuse of a shared L1, a lack of connection between tasks and institutional demands, and managing students' off-task behaviour (Butler, 2011; Carless, 2004, 2008, 2009).

It was argued in Chapter 2 that some of these task implementation difficulties could be resolved through different implementation conditions, but first, an ecologically valid task needed to be used. To find one, an exploratory study was carried out. This study found that the values clarification (VC) task was appropriate for learners' needs in this context because while performing the VC task, students reviewed the (institutionally required) content of their programme and simultaneously pushed their language development for both grammar and lexis more than when performing the other investigated tasks. Therefore, the VC task was implemented as the target of the study in the second phase.

The overall design of the study was as follows. It was a two-phase, mixed-methods study. I implemented teacher-designed tasks into an EFL course for freshmen in South Korea. I was the instructor and the participants were students taking a first-year English for academic purposes (EAP) course intended to prepare them for English-medium instruction at that university. The course followed a sustained-content language teaching approach in which students were expected to master course content and to improve language proficiency (Brinton, 2003; Snow & Brinton, 1997). Following a task-supported approach (Ellis, 2003; Samuda & Bygate, 2008), tasks were integrated with content so that task and content supported each other in an ecologically valid way (Van Lier, 1996, 2000, 2004).

The first phase, a one-semester (16-week) exploratory study, was used to identify which tasks and implementation factors warranted investigation in a quasi-experimental second phase. To maintain ecological validity, the tasks examined were those used in previous

cohorts of the course. The second phase investigated the effect, on task performance, of the following implementation conditions: pre-vs. post-task input processing activity (IPA), repetition (task and procedural), and a pre-task metacognitive awareness (MA) raising activity.

9.2 Summary of the findings

9.2.1 Summary of the exploratory phase (Phase 1)

The exploratory study was conducted from March-June 2012 in two classes of 'Practical English', a course for first-year students who majored in subjects other than English. The exploratory study investigated the process of learning in a variety of different tasks to determine if patterns of interaction and language in these tasks reflected the task designer's intention; and, in what ways different tasks provided different opportunities for language learning.

To maintain the natural ecology of the classroom, I designed tasks that integrated with the course content and implemented them the programme at pedagogically relevant points. For the purposes of task design I used the Willis' (2007; 1996b) framework for task based learning and teaching (TBLT) which involves a sequence of: pre-task lexical work → task (activity) → plan → public report → post-task language focus.

The study focused on six main task types: listing, discussion (sharing experiences and opinions), creative role-play, jigsaw listening, values clarification, and competitive role play. The use of these task types emerged naturally, based on the textbook topics and my prior experience teaching this programme. The tasks were chosen for their pedagogical usefulness, and not for their relevance to theory or previous use in published research.

As the data was qualitative, a discursive process of observe, analyse, reflect, triangulate and interview (participant validation) was followed. The main findings were: (1) the students generally followed the workplan and engaged in extensive English-language collaborative assistance (Ohta, 2001). (2) Students rarely used the 'planning' phase of the task cycle to rehearse their public reports. They used this time to add new content, often using L1. (3) Throughout the task cycle talk about grammar was largely absent except at

times when one of the students was writing down their team's answer. (4) As reported by Carless (2008), students used markedly more L1 when the tasks were more "absorbing" because they worked to create interesting content rather than to develop their L2 (Carless, 2008, p. 335). (5) Of the six task types, the values clarification (VC) task was found to have best prompted students to discuss both lexis and grammar and to do so using relatively little L1.

Other findings included incomplete mining of forms from input (cf. Boston, 2008) lead to incomplete understanding in jigsaw retelling tasks, L1 use increased over three repetitions of the creative simulation tasks, and some students questioned the value of talking in English when doing tasks.

To sum up, the exploratory study was observational and reflective. Its findings raised questions about how students performed the tasks, how they used input, whether or not repetition helped, and the integration of writing and speaking for both lexical and grammatical development. All of which were investigated in the second phase, the intervention study.

9.2.2 Summary of the intervention study using the values clarification task

9.2.2.1 The basis for the interventions

The values clarification (VC) task was chosen for in-depth analysis because it had been found that students talked about both the meaning and the form of the task's prompts because of the VC's task requirement to re-write the consensus-version of the prompt. As the VC task reflected the contents of the related textbook chapter, it allowed students to discuss, and thereby review the contents. Therefore, the VC task connected directly with the goals of the programme because it provided opportunity for meaningful expression, language development and topical review, making it a suitable, ecologically situated task for these students to perform.

The exploratory study showed that learners had a difficult time mining (Boston, 2008) elements from input for use in a retelling procedure, and the missing of key elements led to incorrect retelling. So, consciousness raising (CR) activities that guided learners to identify and process specific lexical items from task input were developed (Fotos, 1994; Rutherford, 1987; D. Willis & Willis, 1996) to determine if learners could use items targeted in this way. These CR activities were made to reflect the Willis' ideas of

focussing on language in the post-task (J. Willis & Willis, 2007; J. Willis, 1996b).

In addition, interviews conducted at the end of the exploratory study indicated that for some participants there was a lack understanding as to why we did tasks and how they would help language learning. So, I decided to use an activity to raise the learners' metacognitive awareness (MA) of the relationship between talk and learning, and how a particular task engaged different, integrated aspects of learning (cf. Ewald, 2004).

Finally, in the exploratory study, I had observed more L1 use when students did the creative simulation task for a second and third time than during the first time. Task repetition had been reported to improve task performance (Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005; White, 2011). Therefore, both types of repetition, task and procedural (Kim, 2013), were investigated further in the next phase.

9.2.2.2 The intervention study

The intervention study took place in the semester immediately after the exploratory study. There were four existing classes of students, two of which were assigned to one condition and two to another. Participants were assigned into constant teams for the VC tasks. The study examined three VC tasks covering six lessons. Treatment conditions were: the presence or absence of a pre-task input processing activity; task repetition and procedural repetition (Kim, 2013); and a pre-task metacognitive awareness (MA) raising activity. The dependent variables were: language related episodes (LREs), the number of turns using L1 (on and off-topic); the number of correctly used forms from the input processing activity; and the correct use of a targeted, 'task natural' phrasal verb (Loschky & Bley-Vroman, 1993).

Input processing activity

To investigate the impact of the input-processing activity (IPA) two conditions were made: input-then-task and task-then-input. In both conditions, students repeated a variation of the same task at the start of the next lesson a few days later (as part of the task repetition study) without input. The design was counterbalanced across two tasks (four lessons) so that all participants were in both conditions. The input-then-task condition allows for examination and familiarisation with language before using it. The task-then-input condition follows Willis' (1996) cycle of pretask, task (task-plan-report), post-task 'language focus' (CR and practice).

The results showed no effect of the input processing activity on task performance. The targeted items from the input were not used at all in any conditions. This was an unpredicted result and qualitative data was used to explain. It was observed from video recordings of teams doing the VC task that the participants did not look at the input scripts once the VC task started, nor did they look at the whiteboard where the items were written. In the metacognitive awareness (MA) intervention and interviews, many participants responded that the input scripts were interesting because they showed the perspectives of native speakers, but they did not consider them as sources of language. When interviewed, participants recommended direct intervention, for example, requiring them to use the items in class or on a test.

Task repetition and procedural repetition

Because there was no result from input in the input-processing study, I collapsed the data and treated the two conditions as equal and investigated whether or not task repetition (doing the same task again) or procedural repetition (following the same task structure with new topical content) had any effect on task performance.

This study involved three tasks across six lessons. As such, both task repetition and procedural repetition were variables (cf. Kim, 2013). Each task had one topic, with two sets of very similar prompts. The first set was done as part of the input-processing study, and the second set of prompts was used at the start of the next lesson (task repetition). The VC task was done in the same groups, allowing for a within-between repeated measures design. Kim (2013) found that generally, language related episodes (LREs) rose with repetitions of tasks, with the exception of lexical LREs in the task repetition condition. This study was predicted to reflect Kim's (2013) result. In addition, L1 use was expected to decrease for task repetition and procedural repetition conditions. For both LREs and L1, the results trended opposite to predictions. There were no trends for the other measures.

When examined qualitatively, it was found that as the number of procedural repetitions increased, and in the task repetition condition, some teams engaged in more off-task talk in L1, and others avoided doing the task by claiming to be finished or by agreeing with all the prompts without discussing them.

The metacognitive awareness (MA) raising intervention changed this behaviour for some

teams. The total number of LREs increased and the number of L1 turns decreased after the MA survey. This change was statistically significantly different from the immediately preceding lesson but not from baseline performance (Lesson 1).

Metacognitive awareness (MA) raising

The metacognitive awareness (MA) intervention was done late in the study. It was a survey in which participants interviewed each other in new groups and discussed ten questions related to the VC task, the teacher's purposes, the way language is learned, and what they do when they do not understand. When they finished the survey, they summarised their discussion and presented it to the class. Finally, the instructor led a plenary question and answer in which he and the participants shared views and expectations about the VC task and language learning in general.

Immediately after the MA survey, participants moved back to their VC teams and did a new VC task. Some teams' VC task performance was dramatically changed by the MA intervention. One team that used a large amount of off-topic L1 used almost no L1 after the MA survey was carried out. Other teams showed little change. In some cases this was due to floor effects – they could not use less L1, because they rarely used it anyway. In other cases, the MA activity did not prompt change. These teams typically reported using L1 because they were talking to friends. These differences in the effectiveness of the MA intervention were explained using Duff's ideas about language socialisation and the norms of classroom behaviour (Duff & Kobayashi, 2010; Duff, 2002, 2010).

9.3 Limitations of the study

As with all research, this study has some limitations to its design that may have influenced the outcomes.

9.3.1 Small number of participants

Due to the situated nature of the intervention study, taking place in existing classrooms, there were only 22 students in Group 1 and 24 in Group 2. They were arranged into seven teams per group. Particularly for between-subjects comparisons, the number of teams was

small, leading to the use of non-parametric statistics, and low statistical power. Recruiting other instructors would have introduced ‘instructor’ as another variable. I compensated for the low numbers by using a repeated measures design to increase the statistical power of some aspects of the study (within-subjects), and by drawing on qualitative data to illuminate the quantitative results.

9.3.2 Measure of L1 use

The measure of L1 use was defined as a turn with any L1 Korean in it, excluding interjections, proper names, and L1-mediated LREs. This is a very broad characterisation. This definition counts all code-switched turns as L1 only, for example, and does not distinguish between one-word and multi-sentence turns. For the VC data set, there were very few of either of these cases. Code switching within a turn was almost invariably in an LRE or an interjection, and L1 use was characterised by quick turns, not by long individual stretches.

I had two main reasons for keeping the turn as the unit rather than use word counts. First, turns represent opportunities to speak, and at each opportunity, each speaker makes a choice about which language to use – therefore turns of L1 represent choices. In this way, the choice to speak in L1 is foregrounded, not the total number of words or time spent. Second, my own ability in Korean (my L2) is only high-beginner/low-intermediate and therefore not strong enough to accurately transcribe all their multi-party talk. Although, I could identify the topic of their L1 talk and confirm it through interviews or with the aid of a native speaker of Korean.

9.3.3 Performance, learning and use

In this study, I used counts of Language Related Episodes (LREs) and counts of ‘target items used in task performance’ as indicators of learning. From SCT this is acceptable, but from cognitive perspectives, this is problematic. LREs have limitations as a measure of learning because, although they are an indication of potential learning taking place, in that they represent learners’ engagement with the language (Swain & Lapkin, 1995, 1998, 2001), they do not necessarily represent complete and accurate knowledge of the item. This was exemplified by Sara’s need to engage in an LREv about the word ‘*stereotype*’ in two different lessons (Extracts 7.12 and 7.13, above). Therefore, from a cognitive perspective, this study should be considered one of task performance, with the assumption,

that the measures of performance are related to learning.

Likewise, the ‘use’ of a new item in task performance is a very high standard to hold students to, particularly if the target items had never been previously encountered. Indeed, some of the students reported that the items were ‘too new’ to be used and the teacher should have them practice or give them a test, before they could use the new items (Extracts 6.6 and 6.8). More sensitive measures, such as vocabulary receptive recall pre- and post-tests (Nation, 2001, pp. 30–31) or the Vocabulary Knowledge Scales (Paribakht & Wesche, 1993) might have detected changes in recognition knowledge that are hidden when items are not used.

However, the design of the study, with multiple task repetitions over a period of several lessons meant that a post-test after one task would influence performance on the next task. It would also dramatically change the nature of the class, from a task-supported method to a test-teach-test format. Furthermore, given the time difference between the first and sixth lesson, a final post-test would be time-biased for items from later lessons. Therefore, pre- and post-tests were not used.

9.3.4 The possibility of response bias

Response bias, or halo effects, are common in interviews because participants pick up on subtle cues given by the interviewer that indicate what the researcher wants them to say, thereby influencing the results (Mackey & Gass, 2005, p. 174). This was a potential issue in this study because of the asymmetrical power relation between me (their teacher) and the participants (students) (Nunan, 1992, p. 150). I tried to minimize this by conducting the interviews after the end of the semester, once all grading had been completed. This reduced my immediate power over the participants, but introduced other problems. By the time of the interviews, some participants had returned to their hometowns outside of Seoul, others had gone on holidays. As a result, I was only able to interview ten participants.

Another factor that might have created a response bias was the presence of the audio recorders on each team’s table. The presence of a recorder may have influenced the results of the study by pushing the participants to use less Korean and stay on task longer than if they were not being recorded, that is, the recorders may have promoted better task performance.

9.4 Theoretical implications

9.4.1 The role of input

The current study was based on Willis' (J. Willis, 1996b) pedagogical framework of TBLT. This framework is rooted in cognitive approaches to language learning and leans heavily on the concept of students noticing a gap between what they know and what they needed (Schmidt, 1990) as a motivation for 'knowledge construction' (Skehan et al., 2012). Willis' argument that the task should come before the language focus was predicated on this idea.

This study investigated the role of noticing in task performance through the placement of the input processing activity (IPA) (Chapter 6). The second time participants did a task, i.e. the repeated task was of particular interest. If students in the task-then-input condition used more target items from the input during the repeated task performance, then we could conclude that the students had noticed the gap during their original task performance and attended to the input items during the post-task language focus. This would support the idea that noticing a gap during task performance drives learning. If students in the input-then-task condition used more target items in the repeated task, we could argue that pre-teaching the items leads to use, and even if noticing the gap did happen, pre-teaching or modelling (cf. Kim & McDonough, 2011) was more effective for promoting item use. However, virtually no use of target items from either task-then-input or input-then-task occurred.

The limitations (Section 9.3.4) pointed out that perhaps more sensitive measures were needed for measuring learning and 'target item use' was too high a standard. However, there the case of Nancy using the expression, '*the range of [sth]*' two lessons after it had been part of the language awareness activity. This delayed use of a target item suggests that noticing did take place, and that moving from noticing items to using them was either difficult or was not pushed by task demands. In order to push item use, the students suggested using traditional approaches such as testing the items and forcing students to use the items during the task. These suggestions indicate that the students felt that more explicit approaches were needed for language learning, and they felt learning was a prerequisite, not a co-requisite for use.

The lack of push from the task was illustrated in the case of ‘*agree with*’. This item was not task essential, but task natural (Loschky & Bley-Vroman, 1993). It was never targeted in the language awareness activities during the study, and it appeared naturally in the transcripts and instructions. However, students could express agreement using: “*I agree with NP*” (correct), “*I agree*” (correct), or “*I agree to NP*” (incorrect) and be understood equally well in all cases. Target-like use of this phrasal verb went unchanged because there was no push from the task demands to use only the correct form. Such a push would involve setting a task such that “*I agree with NP*” and “*I agree to NP*” * would result in different interpretations by the listener. It appears then that to improve the correct use of a form, the correct form should be task essential.

If task essentiality is helpful for noticing, it follows then that seeding the target items into the task input, into the prompts in the case of the VC task, would be a good strategy for task designers to take. In Chapter 5, I reported there was considerable recycling and use of the language in the prompts. The prompts form a task essential part of the task, and the words and forms in them were the focus of most of the task talk. This seeding was not done in the present study, but remains a potential avenue for teachers and researchers to explore.

From a sociocultural perspective, the students did not align with (Atkinson et al., 2007) the IPA as a resource of new knowledge. They reported it was interesting, but ignored it once their task started. One student outright doubted its usefulness for learning new language (Extract 6.11). Perhaps, students’ different alignment with the input than had been predicted influenced the process of noticing, indicating that noticing of language in the native speaker scripts did not happen in the manner proposed by Willis (1996b). However, the task-essential language of the prompts was noticed and attended through LREs. This suggests that noticing coincides with immediate need, not potential need or previous need. Seeding the prompts with target items was recommended for teachers, it might also provide more information on noticing in task performance for researchers. In this way, sociocultural influences offer an explanation for when and how cognitive processes might engage during task performance. This again implies that the ecological approach taken in this study was its strength.

9.4.2 The role of writing

Besides noticing, another underlying construct of the Willis (1996b) framework for TBLT

was the task cycle: activity → plan → report. She claimed that having a public report pushed students to attend to the formal accuracy of their report and thereby improve their output. This was found in the exploratory study not to be the case. In many tasks, planning time was only used to add content through L1 and the presenters for each team would then spontaneously translate whatever they remembered when publicly presenting.

However, in the VC task the final report included rewritten versions of the prompt, which were collected by the instructor. It was reported in Chapters 3 and 5 that while a team member was writing, he or she and the other members would engage in more grammar-related LREs. Therefore, the collaborative re-writing spontaneously pushed noticing of form. This finding supports the work of Williams (2012) who claims that the increased time and the permanence of written forms provide key conditions for a focus on form and the retrieval of explicit knowledge (Williams, 2012, p. 323). She writes,

Surely writing is not a requirement for this [knowledge construction] to occur, but to the extent that the permanent record left by writing increases the demand for attention to formal language features, writing would seem to provide the ideal environment for such co-constructed knowledge (Williams, 2012, p. 325).

This study corroborates her argument and found that the written component need not be long (e.g. a paragraph or an essay) to push attention to form. The prompts in this study were not longer than two sentences, yet grammar-related LREs were found while writing was taking place.

When this finding is viewed through a cognitive lens, the change of mode, from speech to writing, slows down mental processes, and provides a permanent, shared frame of reference to discuss. Speakers no longer have to hold items in short-term memory, because they are permanently stored on the paper. This frees up resources for attention to form. From SCT theory, the students align with their role as learners and the idea that EFL writing should be as correct as they can make it. The written output is treated as if it were any other ‘classwork’ a teacher could evaluate and as the opportunity to develop their grammar (Extract 5.6). Therefore, the alignment shifts from expression to revision during writing, but both alignments fit an overall alignment with their role in the classroom as learners.

Keeping in mind that traditionally, VC tasks did not involve a written component, a

pedagogical implication of this section is that it was not necessarily the public report that pushed a focus to accuracy but the act of collaborating to *write* the final version. The potential difference in the amount of attention to form when a VC task does not include a written revision of the prompt, as in Brinton's (2003) example, versus when they do include such a revision, is a potential avenue of research on the relation of writing to the noticing of form.

9.4.3 The role of repetition

With respect to the findings about task repetition, previous research had predicted increased LREs when tasks were repeated (Kim, 2013) and this study found (non-significant) trends in the opposite direction. In Chapter 7 it was argued that the key difference between this study and Kim's was that she had changed group members for each repetition, whereas, I had kept group membership constant. By controlling group membership, I introduced a potential intervening variable, experience with their partners. This familiarity increased with each repetition. I argued that learners noticed not only language from the input, but they noticed their interlocutor's opinions and ability level. They then used this knowledge when formulating the language and content of their talk in the repeated performances, resulting in fewer LREs. In short, they are engaging with the person, and not only with the task or the language (cf. Platt & Brooks, 2002).

In addition, in a repeated task, students have less need to express their opinions and arguments completely because their interlocutors can remember what was said a few days earlier. The repeated tasks were not greatly different from the original tasks (see Chapter 4 and Appendix 1 for examples), so despite surface differences in the form of the prompts, a similar opinion could be expressed. Without the intellectual challenge of expressing a new opinion, some students became bored and used off topic L1 in greater amounts (e.g. Team A3) and others used task-completion strategies to avoid the task (e.g. Team B3).

These results indicate that task-repetition studies are less straightforward than they appear. The repetition effects reported in other studies (Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005; White, 2011) might be contingent on changing partners, i.e. interlocutor unfamiliarity (cf. Plough & Gass, 1993). In other words, keeping team membership constant created a new, albeit temporary, ecosystem nested within the classroom ecosystem (Bronfenbrenner, 1979). As with any ecosystem, the relationships among the members of that system (here, within the team) and between this ecosystem and the one

in which it is nested (the class) emerge over time (cf. Van Lier, 2004). The nature of the form that emerges depends on the relationships the individual team members have with the higher ecosystems (the class, the university, society, etc.). The opportunities for learning afforded by repeating the task, although recognised individually in interviews (see Extracts 7.1 – 7.4), were not used because these emergent systems within the team aligned team members with other aspects of the context, such as task-completion (e.g. Extract 7.11), task avoidance (e.g. Extract 7.9) and friendship (e.g. Extract 6.2). If team membership were changed each time the task was performed, the rules and relationships within these (very temporary) nested ecosystems would not have formed to such an extent, allowing the overall ‘classroom’ rules and relationships to exert a greater influence on task performance. By considering the ecology of the classroom, it can be claimed that it was more strongly formed interpersonal relationships that inhibited students from taking full advantages of task repetition in the current study.

9.4.4 The role of metacognitive awareness

The metacognitive awareness (MA) intervention had differential impact on teams. In Chapter 8, it was argued that its effect was due to the nature of their peers’ responses during the MA activity. In the case of high off-topic L1 users, when their peers promoted an orientation to the task consistent with the teacher’s beliefs, the off-topic use of L1 was eliminated, and their task performance became longer and more collaborative, that is it had more LREs, more co-construction and more use of continuers (Foster & Ohta, 2005; Ohta, 2001). If the participants’ peers did not promote such an orientation, task performance remained unchanged. With respect to the use of input, students doing the MA activity reported they did not use the input, and there were no cases of peers challenging this and explaining why they should use the input. In the absence of such a challenge from peers, the subsequent VC task target items continued to be ignored, despite the instructor suggesting that they do use the items. This means students adjust their performance to peer-negotiated classroom norms. In other words, in this classroom ecology, keeping one’s behaviour aligned with peer behaviour was more important than the need to align with the teacher. When peer and teacher were aligned, task performance became more in line with task design. When peer and teacher were not aligned, the students in this study aligned with the peer. In this way, the metacognitive awareness raising survey highlighted the differences in task alignment among the various teams and showed how these impacted the statistical data on task performance. It also challenges the

assumption that tasks can mechanically push learning in one direction or another. Learners in this context did not necessarily do what is expected of them by their instructor. They do what is expected of them by their peers. Research in other contexts will need to bear this in mind – it is not only the task that drives learning, but the learners' engagement with it, as mediated by their alignment with the peers in their context.

9.4.5 The approach to research

The previous points raise questions about research on language learning. Although averages, effect sizes and statistically significant differences are important pieces of information, they hide the *events* that contributed to the data. With any average, there are above and below average participants. The ecological approach taken here allowed for the examination of how the numbers arose, and helped identify outliers (e.g. Team A3's use of L1 change after the MA survey). By investigating these outliers, we can get a better understanding of what happens during task performance that was not planned.

More importantly, using students in context who were working toward contextually demanded learning goals under contextually normal constraints highlighted some of the reasons that research is not implemented into classrooms. For example, some teachers might not be able to move students to new partners due to classroom seating arrangements. Task repetition might not be appropriate for these situations, despite evidence that it is helpful (e.g. Kim, 2013). Also, in contexts where peer pressure is high (e.g. working with adolescents), task performance may not be aligned with teacher defined objectives, and further metacognitive work may be necessary for tasks to be successfully implemented.

Therefore, researchers need to account for the ecology of the situation for which their work is intended. In this study, the difference between predictions based on findings from previous, experimental studies and the results with respect to input use and task repetition could be explained only through the voices of the participants. Their voices were sometimes contradictory, but they combine the knowledge of the learner with the knowledge of the researcher and provide a richer and more complete account of task performance. More classroom situated mixed-methods research is needed to examine the interface of theory and practice in EFL contexts.

9.5 Pedagogical contributions

This study makes four main contributions for TBLT pedagogy: research on a previously un-researched task, suggestions for how to integrate target language into task performance, approaches for maximising the benefits of task repetition, and a tool for promoting metacognitive awareness through which task performance can be modified.

First is the usefulness and adaptability of the values clarification (VC) task. Although found in teacher-training books (Brinton, 2003, p. 217; J. Willis & Willis, 2007, p. 76) and in course books themselves (e.g. Sanabria & Sanabria, 2008, p. 119), the VC task has never been researched systematically. Its highly integrated nature and shifting demands: from reading comprehension to expression and discussion of opinions to revising the language of a statement to rewriting the statement to public explanation and justification of the revisions, pushed students through four phases each emphasising different language functions (Chapter 5). In this study, the students understood why I used the VC task and found it to be helpful for practicing giving their opinions, developing their vocabulary and grammar and for revising the course content.

The next suggestion for pedagogy arising from this study pertains to the integration of target language in task performance. Based on the findings reported in Chapter 6, that no target items from the input were used, and the findings reported in Chapter 5, that the task prompts were heavily engaged and recycled, I suggest that teachers include target grammar and vocabulary into the prompts of a VC task. The input processing activity (IPA) featuring a recording and transcript of native speakers doing the same task was found to be interesting, but the target language from it was not used. The delayed use of *'the range of'* by one student suggests that items covered in the input-processing activity might become available for use by learners in the mid- to long-term, but not within that same lesson.

If items cannot be seeded into the task prompts, anecdotal evidence from video recordings suggested that the physical placement of target language items should be such that they are visible during task performance (Chapter 6).

The study also offers some advice to teachers about repeating a task. When the findings reported in Chapter 7 are considered in relation to previous research (e.g. Kim, 2013) I

suggest that changing partners when repeating a task is preferable for pushing language development. In the intervention study, I kept the students in constant teams and they reported that being in the same group made them feel safe and comfortable. This also made them less likely to try new language as they became more comfortable with the members of their team, as evidenced by reductions in LREs across repetitions. So, although it was not directly addressed in this study, it appears that having new partners was a factor involved in the positive results of task repetition reported elsewhere (Kim, 2013; Lynch & Maclean, 2000, 2001; Pinter, 2005).

The study also recommends integrating writing into a speaking task. This study found that students engaged with form (i.e. had more LREg's) while one member was simultaneously writing their final version (i.e. their public report). The values clarification task can be done without revising the statement. However, I argue that the re-writing of the prompt is the driving force behind the attention to form in the VC task. Teachers may wish to find ways to integrate writing into speaking tasks in order to increase the push for more accurate language production (Williams, 2012).

Finally, the MA intervention prompted learners to refocus on why they do tasks and the language-learning potential they represent. The main effect of the MA task pertained to output, increases in LREs and reductions in L1, however it did not impact the use of target items from input. However, the effect the MA task had reinforced the idea that an on-going dialogue with students is needed in order for them to get the maximum benefit from the learning opportunities provided by the task as suggested by Crabbe (2003, 2007).

A last thought with respect to pedagogical contributions is that many teachers work in test-preparation settings. The VC task represents a potential way of improving responses to opinion-oriented questions participants may be required to answer on common large scale proficiency tests. For example two questions on the TOEFL iBT speaking test require test candidates to make an opinion-based choice and then give reasons for their choice (Educational Testing Service, 2013). Because the VC task promotes both expression of opinion and focus to form, it appears that it may be useful for this type of test preparation, too, but this requires further investigation.

9.6 General Conclusions

The general conclusions are listed below.

1. Values clarification tasks provide a platform for expression of opinion and focus on form. Learners recognise the multiple components of the task and can use them for learning.
2. If we wish learners to use input during the task performance, it needs to be an essential part of the task, not simply related to it, or an example of it.
3. Repeating a task does not necessarily help. Learners need to know the benefits they receive from repetition, and the interlocutor needs to be changed for the repetition.
4. The MA survey was one useful way to improve learner knowledge of the purposes of the task, and task performance. However, it was reliant on the answers given by peers. When peers were not aligned with task performance, task performance did not improve.
5. Integrating a collaborative writing component into a speaking task will aid in promoting a focus on form.
6. Researching tasks should include mixed methods design. Statistical data showed trends in classroom patterns that are related to learning. However, it took the examination of transcripts, video and interviews to understand the processes involved in VC task performance.

9.7 Final personal statement

This study investigated tasks in the situation from which their need arose, that is, ecologically valid tasks. Because of my strict adherence to 'pedagogical need' as the primary consideration for task design and implementation, I feel that the greatest strength of this research is its ecological validity. It sought to address issues with tasks that have

been reported by teachers in previous research and issues that I had noticed in my experience as an instructor. It took a teacher-researcher perspective from one context and sought to provide answers that could address other contexts. It was, to cannibalize Lightbown's (2003) title, research in the classroom, for the classroom, done in response to concerns from the classroom.

In this study, I tried to bridge the gap between research and the classroom (Ur, 2012) by situating the study in its ecological setting, and then examining some of the issues in the implementation of tasks and providing ways forward. The values clarification (VC) task was found to consistently push students through four processes, all related to different functions of language: First, to understand the task prompts (comprehension); second, to express their opinions (expression); third, to come to a consensus by adapting the language of a prompt (negotiation); and, throughout the task, to organise their talk and roles (organisation). The task also provided students with the need to focus on both the lexical meaning and grammatical form of the final response: a revised, mutually agreeable prompt. The VC task could be adapted to address many topics and to include pre-specified lexico-grammatical items. It linked speech and writing, which, anecdotally, appeared to assist in directing attention to grammar and push language development.

The study also found that input processing activities did not influence task performance, and suggested practical ways to modify such activities. The study also informed teachers about the effects of task repetition and group membership, that keeping or changing group members will change the learning opportunities from tasks. Finally, the study found that the metacognitive awareness raising activity was helpful in changing some learners' behaviour to one more conducive to language learning, and that this was mediated through the comments of their peers, not the teacher.

My aim was to work within a task-based approach, and not to argue against it, in order to improve the implementation and use of tasks in the messiness of the language classroom. The overall goal of the research was to investigate the implementation of tasks into an Asian EFL classroom in order to address concerns I and others had reported about the use of tasks in Asian contexts (Butler, 2005, 2011; Carless, 2004, 2006, 2007, 2008, 2009; S. Park & Manning, 2012; Swan, 2005). By looking at what students do during tasks and how this changed over time, I was able to address some of these concerns and raise further issues.

LIST OF REFERENCES

- Alegría de la Colina, A., & García Mayo, M. del P. G. (2009). Oral interaction in task-based EFL learning: The use of the L1 as a cognitive tool. *IRAL - International Review of Applied Linguistics in Language Teaching*, 47(3-4).
doi:10.1515/iral.2009.014
- 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.
- Atkinson, D. (2002). Toward a sociocognitive approach to second language acquisition. *The Modern Language Journal*, 86(4), 525–545.
- Atkinson, D. (2010). Sociocognition: what it can mean for second language acquisition. In R. Batstone (Ed.), *Sociocognitive Perspectives on Language Use and Language Learning* (pp. 24–39). Oxford/New York: Oxford University Press.
- Atkinson, D., Churchill, E., Nishino, T., & Okada, H. (2007). Alignment and interaction in a sociocognitive approach to second language acquisition. *The Modern Language Journal*, 91(2), 169–188.
- Batstone, R. (2010). *Sociocognitive Perspectives on Language Use and Language Learning*. Oxford: Oxford University Press.
- Batstone, R. (2012). Language form, task-based language teaching, and the classroom context. *ELT Journal*, 66(4), 459–467. doi:10.1093/elt/ccs058
- Beaumont, M., & Chang, K.-S. (2011). Challenging the traditional/communicative dichotomy. *ELT Journal*, 65(3), 291–299. doi:10.1093/elt/ccq091
- Bialystok, E. (1990). *Communication Strategies: A psychological analysis of second language use*. Oxford: Blackwell Publishers.
- Boston, J. S. (2008). Learner mining of pre-task and task input. *ELT Journal*, 62(1), 66–

- Breen, M. (1987). Learner contribution to task design. In C. N. Candlin & D. F. Murphy (Eds.), *Language Learning Tasks* (Vol. 7). Lancaster and London: Lancaster University.
- Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The Second Language Curriculum* (pp. 187–207). Cambridge: Cambridge University Press.
- Brinton, D. (2003). Content based instruction. In D. Nunan (Ed.), *Practical English Language Teaching* (International Edition., pp. 199–224). New York: McGraw Hill.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by nature and design*. Cambridge, Mass: Harvard University Press.
- Butler, Y. G. (2005). Comparative perspectives towards communicative activities among elementary school teachers in South Korea, Japan and Taiwan. *Language Teaching Research*, 9(4), 423–446. doi:10.1191/1362168805lr176oa
- Butler, Y. G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific Region. *Annual Review of Applied Linguistics*, 31, 36–57. doi:10.1017/S0267190511000122
- 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: Second Language Learning, Teaching and Testing* (pp. 23–48). Essex: Pearson Education Ltd.
- 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 / Philadelphia: John Benjamins.

- Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly*, 38(4), 639–662.
- Carless, D. (2006). Good practices in team teaching in Japan, South Korea and Hong Kong. *System*, 34(3), 341–351. doi:10.1016/j.system.2006.02.001
- Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System*, 35(4), 595–608.
doi:10.1016/j.system.2007.09.003
- Carless, D. (2008). Student use of the mother tongue in the task-based classroom. *ELT Journal*, 62(4), 331–338. doi:10.1093/elt/ccm090
- Carless, D. (2009). Revisiting the TBLT versus P-P-P Debate: Voices from Hong Kong. *Asian Journal of English Language Teaching*, 19, 49–66.
- Cobb, T. (n.d.). Compleat Lexical Tutor v6.2: For data driven language learning on the web. *Compleat Lexical Tutor v6.2: Web VP / BNC-20 v 3.2 British National Corpus lists version (Beta)*. Retrieved August 20, 2012, from <http://www.lex tutor.ca/>
- Coughlan, P., & Duff, P. A. (1994). Same task, different activities: Analysis of an SLA task from an activity theory perspective. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language learning* (pp. 173–194). Westport, CT: Ablex.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Crabbe, D. (2003). The quality of language learning opportunities. *TESOL Quarterly*, 37(1), 9. doi:10.2307/3588464
- Crabbe, D. (2007). Learning opportunities: adding learning value to tasks. *ELT Journal*, 61(2), 117–125. doi:10.1093/elt/ccm004

- Creswell, J. W. (2002). *Research Design: Qualitative, quantitative, and mixed method approaches*. London: SAGE.
- Creswell, J. W. (2007). *Qualitative Inquiry and Research Design: Choosing among five traditions*. Thousand Oaks: Sage.
- Donato, R. (1994). Collective scaffolding in second language acquisition. In *Vygotskian Approaches to Second Language Research*. Westport, CT: Ablex.
- Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: definitions and taxonomies. *Language Learning*, 47(1), 173–210.
doi:10.1111/0023-8333.51997005
- Du Bois, J. W., Corston-Oliver, S., H., Holton, G., & Norris, R. (1998). Voice Walker (Version 2.0). Retrieved from
<http://www.linguistics.ucsb.edu/projects/transcription/tools.html>
- Duff, P. A. (2002). The discursive co-construction of knowledge, identity, and difference: An ethnography of communication in the high school mainstream. *Applied Linguistics*, 22(3), 289–322.
- Duff, P. A. (2010). Language socialization into academic discourse communities. *Annual Review of Applied Linguistics*, 30, 169–192. doi:10.1017/S0267190510000048
- Duff, P. A., & Kobayashi, M. (2010). The intersection of social, cognitive, and cultural processes in language learning: a second language socialization approach. In R. Batstone (Ed.), *Sociocognitive Perspectives on Language Use and Language Meaning* (pp. 75–93). Oxford/ New York: Oxford University Press.
- Educational Testing Service. (2013). TOEFL: Home. Retrieved September 26, 2013, from
<http://www.ets.org/toefl>
- Edwards, C., & Willis, J. (Eds.). (2005). *Teachers Exploring Tasks in English Language Teaching*. Basingstoke / New York: Palgrave Macmillan.

- ELAN (Version 4.6.1). (2013). Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics, The Language Archive,. Retrieved from <http://tla.mpi.nl/tools/tla-tools/elan/>
- Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). *Planning and Task Performance in a Second Language* (Vol. 11). Amsterdam / Philadelphia: John Benjamins.
- Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Quarterly Journal*, 8(3), 19–45.
- 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. (2012). *Language Teaching Research and Language Pedagogy*. Chichester, West Sussex; Malden, MA: Wiley-Blackwell.
- Ellis, R., & Barkhuizen, G. P. (2005). *Analysing Learner Language*. Oxford; New York: Oxford University Press.
- Ewald, J. D. (2004). A classroom forum on small group work: L2 learners see, and change, themselves. *Language Awareness*, 13(3), 163–179.
doi:10.1080/09658410408667093
- 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. doi:10.1080/09658410108667031
- Foster, P., & Ohta, A. S. (2005). Negotiation for meaning and peer assistance in second

- language classrooms. *Applied Linguistics*, 26(3), 402–430.
- Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language Learning*, 59(4), 866–896.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375.
- Fotos, S. S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28(2), 323.
doi:10.2307/3587436
- Gass, S. M., & 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: Lawrence Erlbaum Associates.
- Gass, S. M., Mackey, A., & Ross-Feldman, L. (2011). Task-based interactions in classroom and laboratory settings. *Language Learning*, 61, 189–220.
doi:10.1111/j.1467-9922.2011.00646.x
- Gilabert, R., Baron, J., & Levkina, M. (2011). Manipulating task complexity across task types and modes. In P. Robinson (Ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of language learning and performance* (Vol. 2, pp. 105–138). Amsterdam / Philadelphia: John Benjamins.
- Grice, H. P. (2000). Logic and conversation. In R. J. Stainton (Ed.), *Perspectives in the Philosophy of Language: A concise anthology* (pp. 271–287). Peterborough, Ont.: Broadview.
- Haahr, M. (1998, 2102). RANDOM.ORG - True Random Number Service. Retrieved March 6, 2012, from <http://www.random.org/>

- Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching*, 45(03), 271–308. doi:10.1017/S0261444812000067
- Hawkes, M. L. (2011). Using task repetition to direct learner attention and focus on form. *ELT Journal*. doi:10.1093/elt/ccr059
- Hesse-Biber, S. (2010). *Mixed Methods Research: Merging theory with practice*. New York: Guilford Press.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473.
- Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 330–367. doi:10.1111/lang.12008
- Jeon, M. (2009). Globalization and native English speakers in English programme in Korea (EPIK). *Language, Culture and Curriculum*, 22(3), 231–243. doi:10.1080/07908310903388933
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357–377. doi:10.1093/applin/19.3.357
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of Data-driven Learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (p. 293). Cambridge ; New York, NY, USA: Cambridge 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. doi:10.1111/j.1540-4781.2008.00690.x
- Kim, Y. (2013). Promoting attention to form through task repetition in a Korean EFL context. In K. McDonough & A. Mackey (Eds.), *Second Language Interaction in*

- Diverse Educational Contexts* (pp. 3–24). Amsterdam / Philadelphia: John Benjamins.
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15(2), 183–199.
doi:10.1177/1362168810388711
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition* (1st ed.). Oxford ; New York: Pergamon.
- Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45(3), 261–284. doi:10.1515/iral.2007.012
- Kumaravadelu, B. (1991). Language-learning tasks: teacher intention and learner interpretation. *ELT Journal*, 45(2), 98–108. doi:10.1093/elt/45.2.98
- Kumaravadelu, 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* (pp. 7–31). Newcastle, UK: Cambridge Scholars.
- Lakoff, G., & Johnson, M. (2003). *Metaphors We Live By*. Chicago: Chicago University Press.
- Lam, W. Y. K. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, 13(2), 129–150. doi:10.1177/1362168809103445
- Lantolf, J. P. (Ed.). (2000). *Sociocultural Theory and Second Language Learning*. Oxford [Eng.]; New York: Oxford University Press.
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(01). doi:10.1017/S0272263106060037

- Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. VanPatten & J. Williams (Eds.), *Theories of Second Language Learning: An introduction* (pp. 201–224). Mahwah: Lawrence Erlbaum Associates.
- Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York; London: Routledge.
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8(1), 55–81. doi:10.1191/1362168804lr134oa
- Lightbown, P. (2003). SLA research in the classroom/SLA research for the classroom. *The Language Learning Journal*, 28(1), 4–13. doi:10.1080/0957173038520015
- Littlewood, W. (2004). The task-based approach: some questions and suggestions. *ELT Journal*, 58(4), 319–326. doi:10.1093/elt/58.4.319
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40(03), 243. doi:10.1017/S0261444807004363
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of research on Language Acquisition: Second language acquisition*. (Vol. 2, pp. 413–468). New York: Academic Press.
- Long, M. H. (2007). *Problems in SLA*. New York: Lawrence Erlbaum Associates.
- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. M. Gass (Eds.), *Tasks and Language Learning: Integrating Theory & Practice* (pp. 123–167). Clevedon [England]; Philadelphia: Multilingual Matters.
- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Testing Research*, 4(3), 221–250.
- Lynch, T., & Maclean, J. (2001). “A case of exercising: Effects of immediate task

- repetition on learners' performance. In Martin Bygate, P. Skehan, & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 141–162). Harlow, England: Pearson Education Ltd.
- Macintyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545–562. doi:10.1111/j.1540-4781.1998.tb05543.x
- Mackey, A. (2012). *Input, Interaction, and Corrective Feedback in L2 Learning*. Oxford: Oxford University Press.
- Mackey, A., & Gass, S. M. (2005). *Second Language Research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.
- Mackey, A., & Gass, S. M. (2006). Introduction. *Studies in Second Language Acquisition*, 28, 169–178.
- Mackey, A., Kaganas, A. P., & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41(2), 285–312.
- McDonough, K., & Chaikitmongkol, W. (2007). Teachers' and learners' reactions to a task-based EFL Course in Thailand. *TESOL Quarterly*, 41(1), 107–132.
- Moll, L. C. (1990). *Vygotsky and Education: Instructional implications and applications of sociohistorical psychology*. Cambridge; New York: Cambridge University Press.
- Mori, J. (2002). Task design, plan, and development of talk-in-interaction: An analysis of a small group activity in a Japanese language classroom. *Applied Linguistics*, 23(3), 323–347.
- Mori, J. (2004). Negotiating sequential boundaries and learning opportunities: A case from a Japanese language classroom. *The Modern Language Journal*, 88(4), 536–550. doi:10.1111/j.0026-7902.2004.t01-17-.x

- Murphy, J. M., & Stoller, F. L. (2001). Sustained content teaching an emerging definition. *TESOL Journal*, 10(1), 3–5.
- Nassaji, H., & Jun, T. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, 14(4), 397–419. doi:10.1177/1362168810375364
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge ; New York: Cambridge University Press.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 2–13. doi:10.2167/illt039.0
- Nation, I. S. P. (2009). *Teaching ESL/EFL Reading and Writing*. New York: Routledge.
- Naughton, D. (2006). Cooperative strategy training and oral interaction: Enhancing small group communication in the language classroom. *The Modern Language Journal*, 90(2), 169–184. doi:10.1111/j.1540-4781.2006.00391.x
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17(2), 164–187. doi:10.1177/1362168812460814
- Ngyuen, T. B. T. (2013). *Tasks in Action in Vietnamese EFL High School Classrooms: The role of rehearsal and performance in teaching and learning through tasks*. (Ph.D.). Victoria University of Wellington, Wellington, New Zealand.
- Nunan, D. (1992). *Research Methods in Language Learning*. Cambridge / New York: Cambridge University Press.
- Nunan, D. (2004). *Task-based Language Teaching*. Cambridge, UK ; New York, NY, USA: Cambridge University Press.
- Nunan, D. (2006). Task-based language teaching in the Asia context: Defining “Task.” *Asian EFL Quarterly Journal*, 8(3), 12–18.
- Ohta, A. S. (2001). *Second Language Acquisition Processes in the Classroom Learning*

- Japanese*. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers. Retrieved from
<http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=56339>
- Pally, M. (Ed.). (2000). *Sustained Content Teaching in Academic ESL/EFL: A practical approach*. Boston: Houghton Mifflin.
- Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal/Revue TESL Du Canada*, 11(1), 9–29.
- Park, J.-K. (2009). “English fever” in South Korea: its history and symptoms. *English Today*, 25(01), 50. doi:10.1017/S026607840900008X
- Park, S., & Manning, S. J. (2012). L1 vs. L2 use and classroom interaction in team-taught elementary school classes. *English Teaching*, 67(2), 105–134.
- Philp, J., Oliver, R., & Mackey, A. (2006). The impact of planning time on children’s task-based interactions. *System*, 34(4), 547–565.
doi:10.1016/j.system.2006.08.004
- 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. doi:10.1080/09658416.2010.516831
- 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). Basingstoke / 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, UK:

Cambridge Scholars.

Platt, E., & Brooks, F. B. (2002). Task engagement: A turning point in foreign language development. *Language Learning*, 52(2), 365–400.

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 & Practice* (pp. 35–56). Clevedon [England]; Philadelphia: Multilingual Matters.

Prabhu, N. S. (1987). *Second Language Pedagogy*. Oxford [Oxfordshire]; New York: Oxford University Press.

Richards, K. (2003). *Qualitative inquiry in TESOL*. New York: Palgrave Macmillan.

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 22–57.

Robinson, P. (2007). Criteria for classifying and sequencing pedagogic tasks. In M. del P. García Mayo (Ed.), *Investigating Tasks in Formal Language Learning*. *Multilingual Matters* (pp. 7–26). Clevedon: Multilingual Matters.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Pütz & L. Scola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind* (pp. 239–65). Amsterdam: John Benjamins.

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 / Philadelphia: John Benjamins.

Robinson, P. (2011b). Task-based language learning: A review of issues. *Language Learning*, 61(Supplement 1), 1–36.

- Robinson, P. (2012). Syllabus design. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Oxford, UK: Blackwell Publishing Ltd. Retrieved from <http://doi.wiley.com/10.1002/9781405198431.wbeal1135>
- Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics*, 30(4), 533–554.
- Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45(3), 161–176.
doi:10.1515/iral.2007.007
- Rutherford, W. E. (1987). *Second Language Grammar: Learning and teaching*. London ; New York: Longman.
- Samuda, V., & Bygate, M. (2008). *Tasks in Second Language Learning*. Basingstoke/New York: Palgrave Macmillan.
- Sanabria, K. (2004). *Academic Listening Encounters: Life in Society: Listening Note-taking Discussion* (1st ed.). Cambridge: Cambridge University Press.
- Sanabria, K., & Sanabria, C. (2008). *Academic Listening Encounters: American Studies: listening, note taking, discussion*. Cambridge; New York: Cambridge University Press.
- Sato, M., & Ballinger, S. (2012). Raising language awareness in peer interaction: a cross-context, cross-methodology examination. *Language Awareness*, 21(1-2), 157–179.
doi:10.1080/09658416.2011.639884
- Sato, R. (2010). Reconsidering the effectiveness and suitability of PPP and TBLT in the Japanese EFL classroom. *JALT Journal*, 32(2), 189–200.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied*

- Linguistics*, 11(2), 129–158.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11(1), 11–26.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3–32). Cambridge: Cambridge University Press.
- Seedhouse, P. (2005). “Task” as Research Construct. *Language Learning*, 55(3), 533–570.
- Seedhouse, P., & Almutairi, S. (2009). A holistic approach to task-based interaction. *International Journal of Applied Linguistics*, 19(3), 311–338. doi:10.1111/j.1473-4192.2009.00243.x
- Seifoori, Z., & Vahidi, Z. (2012). The impact of fluency strategy training on Iranian EFL learners’ speech under online planning conditions. *Language Awareness*, 21(1-2), 101–112. doi:10.1080/09658416.2011.639894
- Sheen, R. (2003). Focus on form-a myth in the making? *ELT Journal*, 57(3), 225–233. doi:10.1093/elt/57.3.225
- Shiau, Y., Stacey, & Adams, R. (2011). The effects of increasing reasoning demands on accuracy and complexity in L2 oral production. *University of Sydney Papers in TESOL*, 6, 121–146.
- Shin, S.-K. (2012). “It cannot be done alone”: The socialization of novice English teachers in South Korea. *TESOL Quarterly*, 46(3), 542–567. doi:10.1002/tesq.41
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15(2), 137–158. doi:10.1177/1362168810388692
- Shintani, N. (2012). Repeating input-based tasks with young beginner learners. *RELC Journal*, 43(1), 39–51. doi:10.1177/0033688212439322
- Shintani, N., & Ellis, R. (2011). The incidental acquisition of English plural -s by

- Japanese children in comprehension-based and production-based tasks. *Studies in Second Language Acquisition*, 32(04), 607–637.
- doi:10.1017/S0272263110000288
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing* (pp. 167–185). Harlow, Essex: Pearson Education Limited.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(01), 1–14.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and Lexis. *Applied Linguistics*, 30(4), 510–532.
- doi:10.1093/applin/amp047
- Skehan, P., Bei, X., Li, Q., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, 16(2), 170–187. doi:10.1177/1362168811428414
- Snow, M. A., & Brinton, D. (Eds.). (1997). *The Content-based Classroom: Perspectives on integrating language and content*. White Plains, N.Y.: Longman.
- Song, J. J. (2011). English as an official language in South Korea: Global English or social malady? *Language Problems & Language Planning*, 35(1), 35–55.
- doi:10.1075/lplp.35.1.03son
- Song, J. J. (2012). South Korea: language policy and planning in the making. *Current Issues in Language Planning*, 13(1), 1–68. doi:10.1080/14664208.2012.650322
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158.
- Storch, N. (2008). Metatalk in a pair work activity: Level of engagement and implications

- for language development. *Language Awareness*, 17(2), 95. doi:10.2167/la431.0
- 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. doi:10.1177/1362168810375362
- Storch, N., & Aldosari, A. (2012). Pairing learners in pair work activity. *Language Teaching Research*, 17(1), 31–48. doi:10.1177/1362168812457530
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, Mass.: Newbury.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 124–144). Oxford: Oxford University Press.
- Swain, M. (2006). Languageing, agency, and collaboration in advanced second language proficiency. In H. Byrnes (Ed.), *Advanced Language Learning: The contributions of Halliday and Vygotsky* (pp. 95–108). London: Continuum.
- Swain, M. (2010). “Talking-it-through”: Languageing as a source of learning. In H. Byrnes (Ed.), *Sociocognitive Perspectives on Language Use and Language Meaning* (pp. 112–130). Oxford / New York: Oxford University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371–391. doi:10.1093/applin/16.3.371
- 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. doi:10.1111/j.1540-4781.1998.tb01209.x
- Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the

- first language. *Language Teaching Research*, 4(3), 251–274.
doi:10.1177/136216880000400304
- 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, England: Pearson Education Ltd.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3), 376–401. doi:10.1093/applin/ami013
- Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 58(2), 439–473.
- Tinker Sachs, G. (2007). The challenges of adopting and adapting task-based cooperative teaching and learning in an EFL context. In K. Van den Branden, K. Van Gorp, & M. Verhelst (Eds.), *Tasks in Action: Task-based language education from a classroom-based perspective* (pp. 235–264). Newcastle, UK: Cambridge Scholars.
- Ur, P. (2012, October 16). How useful is Tesol academic research? *the Guardian*. Retrieved September 21, 2013, from <http://www.theguardian.com/education/2012/oct/16/teacher-tesol-academic-research-useful>
- Van Lier, L. (1996). *Interaction in the Language Classroom: Awareness, autonomy, and authenticity*. London: Longman.
- Van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 245–259). Oxford: Oxford University Press.
- Van Lier, L. (2004). *The Ecology and Semiotics of Language Learning: A sociocultural perspective*. Boston ; Dordrecht : Norwell, Mass: Kluwer Academic ; Distributors

- for North America, Central and South America, Kluwer Academic.
- VanPatten, B. (1996). *Input Processing and Grammar Instruction in Second Language Acquisition*. Greenwood Publishing Group.
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52(4), 755–803. doi:10.1111/1467-9922.00203
- VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15(02), 225. doi:10.1017/S0272263100011979
- Vygotsky, L. S. (1986). *Thought and Language*. (A. Kozulin, Ed.) (Revised.). Cambridge, Mass.: MIT Press.
- Wajnryb, R. (1988). The Dicto-gloss method of language teaching: a text-based, communicative approach to grammar. *English Teaching Forum*, 26(3), 35–38.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120. doi:10.1111/j.1467-9922.2012.00729.x
- White, A. (2011). The effectiveness of repeating dictogloss in promoting new vocabulary noticing. *English Teaching*, 66(1), 65–90.
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321–331. doi:10.1016/j.jslw.2012.09.007
- Willis, D., & Willis, J. (1996). Consciousness-raising activities. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language Teaching* (pp. 63–76). Oxford: Macmillan Heinemann English Language Teaching.
- Willis, J. (1996a). A flexible framework for task-based teaching. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language Teaching*. Oxford: Heinemann English Language Teaching.
- Willis, J. (1996b). *A Framework for Task-based Language Teaching*. Harlow, Essex:

Longman.

Willis, J., & Willis, D. (2007). *Doing Task-Based Teaching*. Oxford: Oxford University Press.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006*.

Wray, A., & Fitzpatrick, T. (2008). Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In F. Meunier & S. Granger (Eds.), *Phraseology in Foreign Language Teaching* (pp. 123–148). Amsterdam / Philadelphia: John Benjamins.

APPENDIX 1: SAMPLE VALUES CLARIFICATION TASK

Discussion of Lecture 4: Gender and Language

You have a few minutes to **silently** read and think about the sentences, below. Put a check mark (✓) next to your opinion: strongly agree (SA), agree (A), disagree (D), or strongly disagree (SD). Be prepared to explain the reason you agree or disagree with each sentence. **Do not talk to your partners yet.**

Statement	SD	D	A	SA
(1) It is not a good idea to eliminate gender-biased language because language is too culturally important to change. Reason:	—	—	—	—
(2) Stereotypes about women lead to discrimination against women. Reason:	—	—	—	—
(3) It is more important to have good role models of powerful women on TV and in the real world than it is to change linguistic expressions. Reason:	—	—	—	—
(4) There have been very few real improvements for women, especially working women, over the past twenty years. Reason:	—	—	—	—
(5) There is less gender-bias in Korean language than in English. Reason:	—	—	—	—

Part 2:

When the teacher tells you to begin, talk to your group members about the sentences above. Give your opinions and your reasons and listen to what they say. As you talk, you are to re-write each sentence (use the back of this paper) so that everyone in your group will **strongly agree** (SA) with the new sentence.

Part 3:

The teacher will ask you and the members of your group to go to the front of the room. There, you will have to present your changed sentences and explain why you changed them this way.

Part 4:

The teacher will play a recording of native speakers doing the same activity. Listen to them. Are they the same as or different from you? In what ways?

Prompts used in the repeated version of the VC task

Discussion of Lecture 4: Gender and Language

You have a few minutes to **silently** read and think about the sentences, below. Put a check mark (✓) next to your opinion: strongly agree (SA), agree (A), disagree (D), or strongly disagree (SD). Be prepared to explain the reason you agree or disagree with each sentence. **Do not talk to your partners yet.**

Utterance	SD	D	A	SA
(1) Eliminating gender-biased language will help change the culture in a positive way. <u>Reason:</u>	—	—	—	—
(2) Discrimination against women is caused by negative stereotypes about women. <u>Reason:</u>	—	—	—	—
(3) It is more important to have good role models of powerful women on TV and in the real world than it is to change linguistic expressions. <u>Reason:</u>	—	—	—	—
(4) The life of women at work has improved greatly over the past twenty years. <u>Reason:</u>	—	—	—	—

APPENDIX 2: SAMPLE NATIVE-SPEAKER DIALOG AND LANGUAGE AWARENESS ACTIVITIES

Native-speakers' dialogue transcript

Turn	Speaker	Utterance
1	Dan	So 'stereotypes about women lead to discrimination against women', and I agree because I think that um thoughts you know often lead to words and words lead to action. So if if the words that are being um expressed are- discriminate against women you know that will often lead to uh- real discrimination um against women
2	Gunther	Is your intent to say that it leads to discrimination by men or by other women or by all society?
3	Shaun	I don't think it really mattered I think that you could probably say mm I think here in Korea some of the worst discrimination against women is by other women
4	Gunther	Okay
5	Shaun	Um okay at least the most openly, I don't- I don't know from what I've seen, but uh uh I- I- I have a problem with this uh 'lead to' word there because I think u- the stereotype might be used to justify the discrimination and so the discrimination is leading to the stereotype
6	Dan	Mm-hmm
7	Gunther	Okay
8	Shaun	I don't think it necessarily has to go one then the other I think that that there is a relationship definitely but we often use stereotypes to dehumanize someone or
9	Dan	Mm-hmm
10	Shaun	to justify, like in war, you know the- the- world war two the American army called the Germans ' <i>krauts</i> ' right, which is obviously a racist term, but that's what they did to dehumanize them , so, uh maybe the stereotype is there to
11	Dan	Mm-hmm
12	Shaun	it's I don't know, there's- there's uh what am I saying? There's a back and forth relationship
13	Gunther	Okay
14	Shaun	you know
15	Gunther	symbiotic . okay
16	Dan	You know uh for example um if a person has a stereotype
17	Shaun	Mm-hmm
18	Dan	um about women then oftentimes they'll use language you know that gives um value to that stereotype
19	Shaun	Right
20	Dan	and lan- and if you use um if your language is discriminatory in nature that immediately um is um leads to discrimination. So just there's verbal attack for example
21	Shaun	Mm-hmm
22	Dan	You know if you verbally attack a woman because you have a stereo- a negative stereotype of women. You know just speaking is discriminatory, it doesn't even have to uh speaking and and action co-exist
23	Shaun	Right
24	Dan	So if so, my point is I agree with this statement 'stereotypes about women lead to discrimination against women, both in word and in action.'
25	Gunther	And on the other hand when you look at in life what gets rewarded gets done.
26	Shaun	Right.
27	Gunther	And if you look at say Marilyn Monroe to Paris Hilton
28	Shaun	Mm-hmm
29	Gunther	This behaviour uh society rewards
30	Shaun	Right
31	Gunther	And at certain- a certain point take responsibility for what you're doing because that will be the perception
32	Shaun	Ah I see what you're saying so the- the- Paris Hilton became famous for acting you

		know for (cough) uh I don't know if she deliberately released the sex tape or not, but she made that knowing she was making a sex tape
33	Dan	Mm-hmm
34	Shaun	Out goes the- the- it goes out into the world and poof she's famous. Marilyn Monroe poses for <i>Playboy</i> and gets super famous, so that's the reward //you're talking about, the fame? \\\
35	Gunther	//yeah but but that's a role model \\\ of rewarding a role model stereotype
36	Shaun	Right
37	Gunther	And like Dolly Parton said 'it takes a lot of money to look this cheap' [laughter] You know it takes a lot of brains for Paris Hilton to make that money but i- I see those- those stereotypes as a role model
38	Shaun	Okay
39	Gunther	And until society stops rewarding bimbo blondes
40	Shaun	Mm-hmm
41	Gunther	There will continue to be
42	Shaun	Okay
43	Gunther	this type.
44	Shaun	So you agree with this statement, too?
45	Gunther	Yes
46	Shaun	Um
47	Dan	Do you agree with this statement?
48	Shaun	Well, I I still have that that- my issue was with that words 'lead to'
49	Dan	It might not lead to
50	Shaun	Well, or, discrimination might lead to a new stereotype, I I'm thinking that it could go in the other direction, it could go both ways.
51	Dan	Mm-hmm well they do
52	Shaun	But that doesn't mean this sentence is wrong, if you thing about it. 'Stereotypes against women lead to discrimination against women, AND discrimination against women also leads to stereotypes against women.'
53	Dan	Yes I agree with that too.
54	Gunther	Okay, what about 'stereotypes about women can lead to uh justification for discrimination'?
55	Dan	Well there's no just- there's never justification for discrimination
56	Shaun	Rationalization, but
57	Dan	Mm-hmm
58	Gunther	Rationalization okay yeah
59	Shaun	But I think if we just stick the word AND on the end of that like what I just said 'stereotypes about women lead to discrimination against women and discrimination can- against women can create new stereotypes against women' I think // that would be fine with me\\\
60	Dan	//I agree that's great\\\
61	Gunther	//Okay\\\

Comprehension Questions

1. Which of the three speakers agreed with the original sentence?
2. What does Shaun say about the relationship between discrimination and stereotypes?
3. What example does Shaun give about 'dehumanizing' someone through language?
4. What does Dan say about just speaking?

5. What does Gunther say actresses become ‘bimbo blondes’ to get famous?
6. What is the implication of Gunther’s statement “in life, what gets rewarded gets done” for society? For raising a child?

Language awareness

1. Turn 2: Gunther says, “*is your intent to say*”. Give an equivalent expression.
2. Turn 48: Shaun says, “*my issue was with ____*”. Give an equivalent expression.
3. Turn 12: Shaun says, “*there’s a back and forth relationship*”. Find the word or expression that Gunther uses that means almost the same thing.
4. Turn 34: What word in this turn means ‘*suddenly*’?
5. Turn 55: Dan objects to Gunther using the word ‘*justification*’. What word does Shaun suggest to replace it? How is this word different in meaning than *justification*?
6. Turn 43: What is meant by ‘*this type*’?
7. Turn 34: What expression does Shaun use that means ‘*ah you mean (_)?*’
8. Turn 18: What expression does Dan use that means ‘*shows that something is important to you*’.
9. Turn 50: Shaun uses 2 synonymous expressions. Write them here.
10. Find 3 expressions that use alliteration (words that start with the same sound), assonance (words with the same sound at any place) OR rhyme (words with the same sound at the end). Write them here.
11. Which phrasal verb (2-part verb) is most common in this script?
Write at least 2 synonyms for it here.
Write one new sentence using it.

APPENDIX 3: METACOGNITIVE AWARENESS RAISING ACTIVITY

Questions about Learning English in Class

Part 1: Read these questions and circle the number, 1, 2, 3, or 4 and think about your reasons.

Part 2: Then, ask the members of your group these questions. Remember to ask follow-up questions and have them explain themselves completely. When they ask you, answer them as fully as you can. If there is something you don't know, ask the teacher.

Part 3: Paraphrase your group's discussion and present it to the class. Then collect all these sheets and give them to the teacher.

	Question	NO ← → YES
1.	Do you know why the teacher asks us to give our opinions about the topics from the book? If not, can you try to guess, please? Explain.	1 2 3 4
2.	Do you know why the teacher asks us to try to (strongly) agree on one, new opinion? If not, can you try to guess, please? Explain.	1 2 3 4
3.	Do you know why the teacher makes us talk in English so much in class? If not, can you try to guess, please? Explain.	1 2 3 4
4.	When you read the scripts and do the language awareness activities, do you try to remember new expressions, grammar or words? Explain why or why not.	1 2 3 4
5.	When you hear your partners use an expression, word or grammar structure you don't know, do you try to use it yourself? Explain what you do when you hear or read some new words or grammar.	1 2 3 4
6.	When you are talking in your group in class, are there ever times when you don't know how to say something in English? How do you feel when this happens? What do you do when this happens? What do you think the teacher would want you to do when this happens?	1 2 3 4
7.	Do you sometimes just <u>agree</u> with your partners, even when you really disagree, just so the group can finish the task? Why do you feel you have to do this?	1 2 3 4
8.	Do you sometimes use Korean when you are talking in class? What do you often talk about when you use Korean (the topic, English you don't know, something off-topic, etc.)?	1 2 3 4
9.	Do you think talking a <u>lot</u> in English helps you learn <u>new</u> English? Explain why or why not.	1 2 3 4
10.	What do you do to learn new English grammar, words or expressions while talking in English during class?	

APPENDIX 4: ETHICS SHEETS - ENGLISH

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wananga o te Upoko o te Ika a Maui



INFORMATION SHEET

Researcher: Shaun Manning, Ph.D. Candidate, School of Linguistics and Applied Language Studies, Victoria University of Wellington

Introduction

The aim of the study is to describe and analyze the spoken English used by university students in order to shed light on the process of English language teaching and learning in classroom settings. This study will involve the use of audio and videotape to record students' interactions in English as well as general classroom practices that influence language learning. Conversations, discussions, interviews and other interactions will be recorded, transcribed and analyzed. This analysis is part of the data collection for the researcher's Ph.D. study at Victoria University of Wellington, New Zealand. The project may also result in publication in academic journals and books in which your speech may be quoted, described, and analyzed.

Classroom language use and language learning events will be analysed through observation and by audio and video recordings. It is hoped that the results of this study will be of use to teachers, language learners and educational authorities. The University requires that ethics approval be obtained for research involving human participants. As such you will be asked to give written consent to take part in this study.

Recording (Video and Audio)

Classroom interaction is the focus of this research, so all classroom interaction will be observed and video recorded from a stationary camera placed in a position that will see most of the classroom. The purpose of this video is to record the location and movement of students (e.g. changing of groups) and the positioning of the class instructor. All students will potentially be videotaped at some point during a class; however, to ensure students' (your) privacy, the video recording will be distorted digitally to prevent visual identification. All audio recordings will be made using small, digital recording devices located at each group to record students' talk while performing language learning tasks.

Participation in this study is *voluntary*, should you wish to withdraw from the study at any time you may do so and any of your recorded speech will be deleted from the recordings and from the transcript. Although all members of the class will be observed, only some will be audio recorded for research purposes. However, it is possible that your voice will also be recorded while a recorder is placed near you. If you have indicated on the consent

form that you do not wish to be recorded for research purposes, your voice will be deleted from the recording and your words will be removed from the transcript.

Shortly after each recording has been made there will be an interview conducted with participants, at which point they will have the opportunity to listen to what has been recorded and, if they wish, request that any part of the recording be deleted. They will also be asked to comment on the recording and provide insight into the process of language learning. There will also be the opportunity in further feedback meetings to discuss the results once the project nears completion.

Confidentiality

Upon the completion of the project all recorded data and transcriptions will be securely archived at Victoria University of Wellington, School of Linguistics and Applied Language Studies, to which access is restricted to approved researchers. All data and notes, both written and recorded will only be available to the researcher and research supervisors for the duration of the research and will be securely stored at all times. In any written reports, articles and in the thesis, **all names will be changed** so as to protect the identities of the participants. Any sensitive information will remain strictly confidential and will be scrambled in any written form of the data.

Important Note

Your choice to participate in this study or to not participate in the study has no impact on the grades you will receive for this class. You will not get bonus points for participating and you will not have points deducted for not participating.

Contact Details

If you should have any questions or would like more information, please do not hesitate to contact me at manninshau@myvuw.ac.nz or by phone at 010 6623 8616. If you wish, you may also contact my supervisors:

- Dr. Jonathan Newton by phone at 04 463 5622 or by email at jonathan.newton@vuw.ac.nz
- Dr. Frank Boers by phone at 04 463 6014 or by email at frank.boers@vuw.ac.nz

Or you may contact the School of Linguistics and Applied Language Studies, Victoria University of Wellington, PO Box 600 Wellington 6140, New Zealand, phone 04 463 5600, email lals@vuw.ac.nz.



CONSENT TO PARTICIPATE IN RESEARCH

The aim of the study is to describe and analyze the spoken English used by university students. This study will involve the use of observation, audio and videotape to record students' interactions in English as they learn in class. Conversations, discussions, interviews and other interactions will be recorded, transcribed and analysed.

The researcher has discussed the project with me and I have had the opportunity to ask him questions about it. I understand that I may withdraw myself (or any information I have provided) from this project (before data collection and analysis is complete) without having to give reasons and *without penalty of any sort*. This means that if I choose to participate I will not get any bonus for my grade, and if I choose to not participate my grade will not suffer. The project may result in publication in academic journals and books in which my speech may be quoted, described, and analyzed.

I understand that my *privacy will be protected* in the following ways. My name will be changed as will the names of others whom I talk to or talk about. Any other part of my speech that may identify me as the speaker will be altered to disable such identification. Recordings and transcripts of my speech will be locked in a safe place or never left unattended. Video recordings will be digitally distorted, (i.e. either pixelated or blurred) to such an extent that visual identification of persons on them will be impossible. In addition, I understand that if I wish for certain things I might say to remain private, I will inform the researcher of this and these things will be deleted from the recording and the transcript.

If I have provided other materials relating to myself, my work, or my speech, or if other materials are publicly available, these materials may also be quoted, described, and analyzed.

By signing this form, I certify that research project has been satisfactorily explained to me and that I consent to participate in it in the ways described above. I also certify that I know that I may choose to discontinue the research at any time, and *without penalty*. I also certify that I understand that my choice to participate or to not participate has no impact on the grades I will receive for the class.

I agree to take part in this research (circle one): **YES** **NO**

Name: _____ Signature: _____ Date: _____

Please tick here if you wish to receive a copy of a short summary of the study: ☐

Mail or email address to which the summary can be sent:



APPENDIX 5: STIMULATED RECALL INTERVIEW QUESTIONS

Procedure: The interviewer will play a recording of a short section of the classroom activity in which the interviewee is one of the participants one or more times to allow the interviewee to hear it clearly. The interviewer will tell the interviewee which day the recording was made. The following questions will be asked.

1. Can you describe what happened as you were doing this task?
2. What were you thinking and feeling as you did the task?
3. Were there any points at which you didn't understand exactly what one of your group members was saying? What did you do?
4. How well do you think you performed the task?
5. What specific aspects of your performance are you satisfied/unsatisfied with?
Why?
6. Did you understand the teacher's instructions?
7. Did you have enough time to do the task well?
8. If you did the task again, what would you do differently? Why?
9. Do you think that doing this task helped your English improve? Explain.
10. Is there anything else you'd like to tell me about this task?

APPENDIX 6: ETHICS SHEET - KOREAN

VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wananga o te Upoko o te Ika a Maui



정보지

연구원: 산 메닝, 박사 학위 과정, 언어학과 응용 언어 연구, 웰링턴 빅토리아 대학교

서론

이 연구의 목표는 교육현장에서의 영어 가르침과 배움의 과정을 해명하기 위해 대학생들이 사용하는 구어체 영어를 묘사하고 분석하는 것에 있다. 이 연구는 학생들의 영어 상호작용뿐만 아니라 어학 습득에 영향을 주는 일반적인 실지수업을 기록하기 위해 오디오와 비디오의 사용을 포함한다. 대화, 토론, 인터뷰를 포함한 그 외의 상호작용들은 녹음, 기록, 분석될 것이다. 이 분석 연구는 뉴질랜드에 위치한 웰링턴 빅토리아 대학의 박사 학위 연구 중인 연구원의 자료 수집의 일부이다. 또한 이 프로젝트는 당신의 말들이 인용, 묘사, 분석되어 학술지나 서적의 형태로 출판되어질 수 있다.

관찰과 오디오, 비디오테이프에 녹화된 것들을 통해 수업 언어 사용과 언어 학습 이벤트들을 분석할 것이다. 이 연구가 선생님들, 언어 학습자들과 교육 당국에 이용될 수 있길 바란다. 대학은 참가자들이 수반된 이 조사에 윤리적 승인이 존재하기를 요구한다. 그렇기에 당신은 이 연구에 참가한다는 것에 동의한다는 문서를 작성하기를 부탁받을 것이다.

녹화(비디오와 오디오)

수업 상호작용은 이 조사의 초점이기 때문에 모든 수업 상호작용은 관찰되어지고, 교실의 전반적인 모습이 잡히는 곳에 위치한 정지된 카메라로 녹화되어질 것이다. 이 비디오는 학생들의 위치와 움직임(그룹 변화) 그리고 강사의 위치를 녹화하기 위한 목적에서 사용된다. 모든 학생들은 잠재적으로 수업시간 중에 녹화되어 질 것이다. 하지만 학생들의 사생활권 을 보장하고 시각적 신원 식별방지를 위해 디지털 방식으로 왜곡되어 촬영될 것이다. 모든 음성 녹음은 언어 습득 과제를 수행하는 동안의 학생들의 대화를 녹음하기 위해 각 그룹마다 위치한 작은 디지털 녹음기로 녹음될 것이다.

이 연구의 참여는 자발적이고, 연구에서 철회하고 싶다면 언제든지 가능하며, 이전에 녹음되었던 목소리와 기록들은 모두 삭제될 것이다. 교실의 모든 구성원들이 관찰되어지지만, 그 중 일부만이 조사 목적으로 음성 녹음 될 것이다. 하지만 당신 근처에 녹음기가 위치했을 경우, 당신의 목소리 또한 녹음될 가능성은 있다. 당신이 조사 목적으로 녹음되지 않길 원한다는 점이 동의서에 명시되면 당신의 목소리와 말들은 녹음과 기록에서 모두 삭제 될 것이다.

각 녹음이 끝난 직후에는 참가자들과의 인터뷰를 시행할 것이며, 무엇이 녹음됐는지 들어볼 기회가 주어질 것이다. 또한 그들이 원한다면 녹음된 부분 중 어떤 부분이라도 삭제하도록 요구할 수 있다. 훗날 프로젝트가 거의 끝나가기 전에 진행될 피드백 미팅 시에도 같은 기회가 제공될 것이다.

기밀성

프로젝트가 끝날 때까지 모든 녹화된 데이터와 기록들은 승인된 연구원만이 접근할 수 있도록 웰링턴에 있는 빅토리아 대학, 언어학과 응용 언어 연구학과에 안전하게 보관될 것이다. 쓰여지고 녹화된 모든 데이터와 기록들은 조사 기간 동안 해당 연구원과 연구 지도교수만이 이용 가능하며 항상 안전하게 보관될 것이다. 그 어떤 보고서, 기사 그리고 논문에서도 참여자들의 신원을 보호하기 위해 **모든 이름은 변경될 것이다**. 어떤 민감한 정보도 엄격한 비밀이며, 모든 문서 형식의 데이터에서 변환되어질 것이다.

중요 사항

이 연구에의 참여여부는 수업에서 받게될 성적에는 아무런 영향을 주지 않습니다. 참여한다고 해서 가산점이 주어지지 않으며, 불참시에도 감점되지 않을 것입니다.

연락처

질문이 있거나 더 많은 정보를 원한다면 manninshau@vuw.ac.nz 혹은 전화로 010 6623 8616 로 바로 연락하시기 바랍니다.

만약 원한다면 본인의 지도교수께 연락할 수 있습니다 :

- 조나단 뉴튼 박사 : 전화 04 463 5622 이메일 jonathan.newton@vuw.ac.nz
- 프랭크 보어스 박사 : 전화 04463 6014 이메일 frank.boers@vuw.ac.nz

혹은 웰링턴에 있는 빅토리아 대학, 언어학과 응용 언어 연구학과, 사서함 600 웰링턴 6014, 전화 04 463 5600, 이메일 lals@vuw.ac.nz. 로 연락주시기 바랍니다.

APPENDIX 7: RECALL INTERVIEW QUESTIONS (KOREAN)

조사 참여 동의서

이 연구의 목적은 대학생들이 사용하는 구어체 영어를 서술하고 분석하기 위함에 있다. 이 연구는 수업 중에 이루어지는 학생들의 영어 상호작용을 관찰하고, 오디오와 비디오를 사용하여 기록할 것이다. 대화, 토론, 인터뷰를 포함한 그 외의 상호작용들은 녹화, 기록되고 분석되어질 것이다.

이 연구원은 이 프로젝트에 관해 본인과 상의 하였으며, 이 프로젝트에 관해 물어볼 기회가 주어졌었다. 본인은 철회 이유를 제공하거나 그 어떠한 불이익 없이 자신(혹은 본인이 제공한 어떤 정보)을 (데이터 수집과 분석이 끝나기 전에) 이 프로젝트로부터 철회할 수 있다는 것을 이해한다. 이 프로젝트는 본인의 말들이 인용, 묘사, 분석되어 학술지에 출판되어 질 수 있다.

본인은 본인의 사생활권이 다음에 나오는 방식대로 보호될 것이라는 것을 이해한다. 본인의 이름을 포함하여 본인이 말하고 있는 대화 상대 혹은 대화에 거론되는 이름들은 변경될 것이다. 본인의 말에서 신원을 알아낼 수 있는 어떤 부분도 신원식별이 불가능하도록 수정될 것이다. 녹음되고 기록된 본인의 말들은 안전한 곳에 잠겨져 보관될 것이며 절대로 지켜볼 사람없이 남겨져 있지 않을 것이다. 비디오 녹화는 디지털 방식으로 왜곡되어 (픽셀화 혹은 흐릿하게 처리되어) 시각적 신원 식별이 불가능할 것이다. 게다가 본인은 만일 어떤 부분이 사적으로 보호받기를 원한다면 연구원에게 이것을 알리고 녹음과 기록으로부터 그 부분을 삭제할 수 있다는 것을 이해한다.

만일 본인의 직업, 연설 혹은 공적으로 사용가능한 본인과 연계된 다른 정보들을 본인이 제공했다면 그 정보들도 인용, 묘사, 분석되어질 수 있다.

이 동의서에 서명함으로써 이 조사 프로젝트가 본인에게 충분히 설명되었고 이 프로젝트에 참여하는 것에 동의한다는 사실을 증명한다. 본인은 또한 언제든지 어떠한 불이익없이 이 조사에 대한 참여를 중단할 수 있다는 것을 안다는 사실을 증명한다. 이 동의서에 서명함으로써 이 조사 프로젝트가 본인에게 충분히 설명되었고 이 프로젝트에 참여하는 것에 동의한다는 사실을 증명한다. 본인은 또한 언제든지 어떠한 불이익없이 이 조사에 대한 참여를 중단할 수 있다는 것을 안다는 사실을 증명한다. 또한 본인의 참여여부는 이 수업에서 받게 될 점수에 어떠한 영향도 끼치지 않음을 안다는 사실을 증명한다.

본인은 이 조사에 참여하는 것에 동의합니다. (동그라미 치시오): **네** **아니오**

이름: _____ 서명: _____ 날짜: _____

이 연구의 요약된 복사본을 받고 싶다면 이곳을 클릭하세요: ☐

요약본을 받으실 수 있는 주소나 이메일을 적어주세요:

APPENDIX 8: TASKS INVESTIGATED IN THE EXPLORATORY STUDY

The six tasks investigated in the exploratory study are outlined here using Ellis' (2003) framework:

1. Goal: the general purpose of the task
2. Input: The verbal or non-verbal information supplied by the task
3. Conditions: The way in which the information is presented or it is to be used (split/shared), (Converging/diverging)
4. Procedures: The methodological procedures to be followed in performing the task
5. Predicted outcomes:
 - a. Product: The 'product' that results from completing the task. It can be 'open' (allow for several possibilities) or 'closed' (allow for only one solution)
 - b. Process: The linguistic and cognitive processes the task is hypothesized to generate. (Ellis, 2003, p.21)

Each description of a task represents a bare-bones version of the task-as-workplan, i.e. what the task designer expected to happen in each task, much of which did not happen as anticipated.

Features of the listing task

Feature	Description
Goal	This task served as the pre-task for a subsequent 'captains & ambassadors' task, so the goal was to generate many potential ways to improve their health
Input	Input was solely from other members of the group; students could refer to their textbook for ideas if they wished
Conditions	The information is shared among members and a consensus is to be reached on items to include or reject (converging)
Procedures	Students brainstorm a list in groups of 3-5 Present their list by reading it aloud to the class.
Outcomes: Product	A list of 20 actions one could take to improve their health. One member reads the list to the class (public report) [Open outcome]
Outcomes: Process	Students will need to generate vocabulary for some items (through LREs, L1, or by referring to their textbook). A minimum interaction would be one in which one student nominates an item and another writes it down in English. It is expected that for some items a discussion as to what the nominator means, or the appropriacy of including it in the list and its possible rejection and argument may happen. [convergent process]

Features of the discussion/sharing personal opinion and experience task

Feature	Description
Goal	To engage learners' ideas and interest in the topic by asking about their views on the matter
Input	The list of questions to talk about, the responses of their group members
Conditions	Shared information; responses and talk can either converge or diverge, however, there may be social pressure to converge.
Procedures	Students silently read a short list of questions and think about their answers. After the teacher tells them to begin they start talking. After a period of time spent talking the teacher tells them to prepare their report. The report is presented either to another group or to the whole class.
Outcomes: Product	A report of either a summary of their talk, or an extended answer to one question from the list. (Reports were assigned to be between 1 and 2 minutes long). [open product]
Outcomes: Process	Students will generate, explain, question and reject opinions. The talk should contain the basic features of everyday conversation – getting and holding the floor, interruptions, repetitions, false starts, etc.; there may be some LREs and some L1; there is a danger that students will go off-topic (and if they do, there may be a great deal of L1) [divergent opinion gap process + convergent summarization process]

Features of the create a role play task

Feature	Description
Goal	To personalize textbook material to students' own context and to consider issues related to changing society
Input	Listening activities that were based on interviews were done before the task (e.g. Sanabria & Sanabria, 2008. p. 4-8), but these are not direct models for the task. Other group members' talk was the primary input
Conditions	Shared information; working toward a common performance
Procedures	After listening to the interview and answering the comprehension questions in the textbook students are assigned into groups and told to imagine one of them is the interviewer and the other members interviewees, and they are to conduct an interview on the assigned topic (related to but not the same as the textbook). Students create the roles, questions and responses. The role play is acted to the entire class
Outcomes: Product	A role play of an interviewer and 2 or three interviewees [open product]
Outcomes: Process	Students work together to brainstorm their roles (what kind of interview, are the interviewees known to each other, do the interviewees disagree, etc.) and the questions they will ask. They then practice asking and answering the questions, working together to improve its spontaneity. As the questions may be written down during their creative phase, LREs on grammar and lexis are anticipated. [convergent creative process]

Features of the jigsaw listening task

Feature	Description
Goal	To learn the lecture material (content) by retelling it
Input	The audio recording of the lecture material. Their partner's notes on the other portion of the lecture.
Conditions	The information is split between 2 partners. The talk should be converging to a mutually acceptable version
Procedures	Students are divided into pairs. Half listen to first part of the lecture while the others wait in the hall. While listening, they take notes about the content of the lecture. They switch roles. After both parts have been played, pairs regroup and attempt to reconstruct the lecture. They collaborate with another pair to make the final version to be reported to the class.
Outcomes: Product	A spoken version of the lecture that contains (most of) the information from the lecture. [closed product]
Outcomes: Process	Students need to listen and take notes onto a blank sheet of paper. They will use the notes to help them reconstruct what they heard to their partner. The bulk of the talk will be content-related, with one person primarily giving information and the other listening/questioning and writing it down. Then their roles reverse. It is expected that students mine the input they listened to in order to rebuild the lecture. Both lexical and grammatical LREs are expected. [convergent process]

Features of the Simulation task

Feature	Description
Goal	To create and successfully 'sell' an idea or concept to another team
Input	Input from the previously heard lecture is not directly related to this task but may be helpful. Input from group members differs at different stages.
Conditions	Shared information, followed by split information.
Procedures	Students in groups of 3 or 4 create their 'product' Students choose roles 'captain, secretary, and 2 ambassadors' Ambassadors move to separate groups Captains and secretaries remain. Ambassadors present to the new group and try to convince the captain and secretary to choose their idea/product; they answer questions and defend their ideas. Captains and secretaries choose ONE winner and justify it in their report to the class
Outcomes: Product	There are 2 products: the initial idea (in the example given, the presentation of the superhero) [open product]; and the report of the final choice made by the captain to the class [open product]
Outcomes: Process	Two distinct processes are expected. A creative-collaborative process in coming up with and developing their idea. [convergent process] An interactive-competitive process of trying to sell a product to another group while challenging the other one. As the process is almost entirely meaning-based, there will be primarily lexically related LREs. [divergent process for the ambassadors; convergent process for the captain/secretary who must choose one winner]

Features of the values clarification task

Feature	Description
Goal	To present a mutually agreed set of statements on a controversial topic
Input	The textbook content provides relevant background knowledge. Other input is the prompt itself (written statement) and their partner's talk.
Conditions	Shared information; convergent talk
Procedures	Students silently read the prompts, and decide their level of agreement and jot down their reasons. They discuss the items and explain their reasons. Where there is disagreement they work together to rephrase the prompt into a version that is agreeable to all members
Outcomes: Product	The revised sentences [open product] Report to the class about revised sentences and reasons for the revisions [open product]
Outcomes: Process	Discussion of the prompt's meaning and their reasons for agreeing or disagreeing should promote lexical LREs. Discussion of the revised version should also promote grammatical LREs. [convergent process]

APPENDIX 9: PHASE 2 DATA

Lesson 1											
			LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Grammar Fo	Lexis Focus	Pronunciati	Total LREs	Correct	Incorrect	Use of targ	L1 on	L1 off
A1	1	14	4	4	0	8	0	2	0	2	2
A2	1	14	2	4	0	6	2	1	0	3	1
A3	1	14	2	7	0	9	0	1	0	55	2
A4	1	14	0	2	1	3	0	1	0	3	0
A5	1	14	4	11	0	15	0	0	0	5	2
B1	2	14	4	1	1	6	0	1	0	3	0
B2	2	14	4	9	0	13	0	4	0	2	1
B3	2	14	7	9	4	20	0	1	1	3	0
C1	2	14	2	2	0	4	0	0	0	0	0
C2	2	14	2	5	0	7	2	0	0	5	1
C3	2	14	1	5	0	6	1	0	0	13	5
C4	2	14	0	7	0	7	1	0	0	2	0
D1	1	11	0	2	2	4	0	1	0	14	0
D2	1	11	2	6	1	9	1	0	0	7	0

Lesson 2											
			LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Grammar Fc	Lexis Focus	Pronunciati	Total LREs	Correct	incorrect	Use of targ	L1 on	L1 off
A1	1	13	0	2	0	2	1	2	0	5	0
A2	1	13	0	5	0	5	3	0	0	13	48
A3	1	13	3	12	0	15	0	3	0	11	0
A4	1	13	1	2	0	3	0	0	0	0	0
A5	1	13	2	2	0	4	0	0	0	0	0
B1	2	13	2	4	0	6	0	0	0	2	12
B2	2	13	1	7	0	8	1	1	0	0	0
B3	2	13	5	4	0	9	0	0	0	20	1
C1	2	10	0	2	0	2	0	0	0	2	0
C2	2	10	0	1	0	1	5	1	0	0	0
C3	2	10	3	5	0	8	0	1	0	16	0
C4	2	10	0	1	0	1	1	0	0	0	3
D1	1	11	3	2	0	5	2	0	0	18	0
D2	1	11	7	2	1	10	0	0	0	2	1

Lesson 3											
			LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Grammar Fc	Lexis Focus	Pronunciati	Total LREs	Correct	Incorrect	Use of targ	L1 on	L1 off
A1	2	14	1	2	0	3	0	3	0	0	1
A2	2	14	0	4	0	4	0	0	0	10	2
A3	2	14	5	3	2	10	2	0	0	17	20
A4	2	14	2	1	0	3	0	0	0	1	15
A5	2	14	1	4	2	7	0	0	0	3	1
B1	1	14	2	3	2	7	0	1	0	3	0
B2	1	14	2	2	0	4	0	1	0	0	0
B3	1	14	5	5	2	12	0	0	0	19	0
C1	1	16	4	5	0	9	0	0	0	0	0
C2	1	16	1	4	1	6	7	0	0	4	20
C3	1	16	1	5	1	7	1	0	0	11	0
C4	1	16	0	5	1	6	0	0	0	1	0
D1	2	14	4	6	0	10	0	0	0	18	0
D2	2	14	4	4	1	9	5	0	0	9	1

Lesson 4											
			LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Grammar Fo	Lexis Focus	Pronunciati	Total LREs	Correct	incorrect	Use of targ	L1 on	L1 off
A1	2	11	0	3	0	3	0	0	0	0	0
A2	2	11	0	5	0	5	2	0	0	3	5
A3	2	11	2	5	0	7	0	0	0	9	19
A4	2	11	0	3	0	3	0	0	0	5	0
A5	2	11	2	2	2	6	0	1	0	0	0
B1	1	10	0	4	0	4	2	0	0	0	16
B2	1	10	1	5	0	6	0	1	0	0	0
B3	1	10	4	2	1	7	4	4	0	15	1
C1	1	8	0	1	0	1	0	0	0	4	0
C2	1	8	0	1	0	1	3	0	0	5	18
C3	1	8	0	3	0	3	1	0	0	16	0
C4	1	8	1	1	0	2	3	1	0	1	0
D1	2	8	2	3	0	5	1	0	0	4	1
D2	2	8	1	4	1	6	0	0	0	7	1

Lesson 5												
				LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Metacog co	Grammar Fo	Lexis Focus	Pronunciati	Total LREs	Correct	Incorrect	Use of targ	L1 on	L1 off
A1	1	16	MA+	1	8	0	9	1	3	0	2	4
A2	1	16	MA+	3	9	0	12	2	0	0	19	1
A3	1	16	MA+	7	12	3	22	0	0	0	0	0
A4	1	16	MA+	5	5	1	11	0	0	0	3	11
A5	1	16	MA+	8	17	0	25	11	1	0	0	0
B1	2	14	MA-	3	2	0	5	2	0	0	1	24
B2	2	14	MA-	5	4	1	10	0	4	0	0	2
B3	2	14	MA-	2	10	1	13	0	0	0	2	3
C1	1	15	MA-	8	3	1	12	1	1	0	0	3
C2	1	15	MA-	0	7	0	7	0	3	0	4	5
C3	1	15	MA-	3	2	0	5	4	1	0	7	0
C4	1	15	MA-	1	1	0	2	0	1	0	2	0
D1	2	13	MA+	5	2	1	8	4	0	0	0	3
D2	2	13	MA+	3	2	0	5	0	0	0	1	5

Lesson 6												
				LREs				Agree + prep			L1 use	
Team	Input order	Time on tas	Metacog co	Grammar Fo	Lexis Focus	Pronunciati	Total LREs	Correct	Incorrect	Use of targ	L1 on	L1 off
A1	1	13	MA+	2	6	2	10	1	0	0	1	1
A2	1	13	MA+	2	3	0	5	0	0	0	14	0
A3	1	13	MA+	6	12	1	19	0	0	0	3	0
A4	1	13	MA+	3	4	0	7	0	0	0	0	0
A5	1	13	MA+	4	6	0	10	4	1	0	3	2
B1	2	12	MA-	2	4	0	6	0	0	0	5	6
B2	2	12	MA-	5	6	0	11	0	3	0	2	2
B3	2	12	MA-	2	2	1	5	0	0	0	6	8
C1	1	11	MA-	4	3	2	9	2	0	0	2	0
C2	1	11	MA-	1	3	0	4	1	1	0	6	1
C3	1	11	MA-	0	4	0	4	1	1	0	2	0
C4	1	11	MA-	2	1	0	3	0	0	0	0	0
D1	2	10	MA+	2	0	1	3	0	0	0	0	2
D2	2	10	MA+	2	3	0	5	0	0	0	0	1

APPENDIX 10: TARGET ITEMS FOR VC TASKS

This appendix contains the list of target language items from each Language Awareness activity for the tasks in this study.

VC ₁	VC ₂	VC ₃
A shop teacher,	Is your intent to say (something)	(sth) has its limitations
Offsetting (= counterbalancing/ balancing out / evening out)	My issue was with ... / my problem was with ... / I have a problem with ...	Pick (sth) apart Hit (= talk about)
On one level, on another level, in one way/ in another way, from a different perspective/point of view	There's a back and forth relationship / symbiotic	Out of context
Beyond the scope / range of (what is possible)	Poof (= suddenly)	Make a case (= convince someone of something)
Historically,	I see what you're saying	One-sided
the best of the best/ the cream of the crop,	Rationalization	Outweigh
	To value (something)	Put something (of your own) into (something) (= make an effort)
	Lead to (this was in the prompt as well, so it was not considered in the analysis)	Have a dialogue with (someone)