

**ENGAGEMENT WITH LANGUAGE DURING TRANSCRIPT
REVISION: JAPANESE UNIVERSITY ENGLISH LEARNERS'
PROCESSES, PRODUCTS, AND PERSPECTIVES**

**Thesis submitted for the degree of
Doctor of Education
in
Linguistics and TESOL
at the University of Leicester**

**by
Jeremy Scott Boston**

**School of Education
University of Leicester**

April 2018

Engagement with Language during Transcript Revision: Japanese University English Learners' Processes, Products, and Perspectives

Jeremy Scott Boston

Abstract

This classroom-based study investigated the use of Transcript Revision Tasks (TRTs) as a means of focusing learner attention to divergences between their output and English norms. TRTs entailed learners self-transcribing recordings of their oral production and then revising transcripts for accuracy in pairs. Previous TRT studies have predominantly focused on the products of transcript revision but understudied how learners arrived at their revisions (their processes). To investigate in more depth how learners attended to form during transcript revision, this study employed Svalberg's (2009, 2012) 'Engagement with language' (EWL) construct which encompasses learners' cognitive as well as social and affective engagement. Data were gathered from four pairs of Japanese university students in an intact English speaking and listening course who performed three TRTs over the course of 15 weeks (one semester). Learners' pair dialogues during transcript revision, their revised transcripts, and end of semester interview protocols were analysed to establish degree of EWL. Subsequent repeated individual performances of oral production tasks after each TRT were analysed for evidence of retention of revision. In-line with prior studies of learner metatalk, this study found more extensive discussion of linguistic choices to be more facilitative of retention than discussions which were perfunctory. This study found that the learners' degree of cognitive engagement (attention to and discussion of language form) correlated with their degree of social engagement (supportive/collaborative interaction) and affective engagement (positive evaluation of TRTs or task partners), and that lack of social and affective engagement deterred cognitive engagement with language forms. Learner retention of revisions was found to be related to the degree of cognitive EWL. This study concludes with discussion how the TRTs could be better designed and implemented to generate learner engagement with language.

TABLE OF CONTENTS

1	INTRODUCTION.....	1
1.1	SUPPLEMENTING COMMUNICATIVE TASKS WITH FOCUS ON FORM ACTIVITIES	1
1.2	TRANSCRIPT REVISION TASKS.....	2
1.3	THE PURPOSE OF TRTs	3
1.4	LIMITATIONS OF PRIOR INVESTIGATIONS INTO CLASSROOM USE OF TRTs	4
1.5	PURPOSE OF STUDY AND RESEARCH QUESTIONS	5
1.6	ORGANIZATION OF THE THESIS	5
2	CHAPTER 2: CONCEPTUAL BACKGROUND	7
2.1	NOTICING	7
2.1.1	<i>The Noticing Hypothesis</i>	<i>8</i>
2.1.2	<i>Noticing and awareness as facilitators of language acquisition.....</i>	<i>8</i>
2.1.3	<i>The Output Hypothesis and noticing</i>	<i>10</i>
2.2	PEER-PEER DISCUSSION OF LANGUAGE	11
2.2.1	<i>Sociocultural theory.....</i>	<i>11</i>
2.2.2	<i>The Zone of Proximal Development.....</i>	<i>12</i>
2.2.3	<i>Collaborative Dialogue</i>	<i>13</i>
2.2.4	<i>Language related episodes.....</i>	<i>14</i>
2.3	ENGAGEMENT WITH LANGUAGE	16
2.3.1	<i>Cognitive engagement.....</i>	<i>17</i>
2.3.2	<i>Social Engagement</i>	<i>20</i>
2.3.3	<i>Affective Engagement</i>	<i>22</i>
2.4	CHAPTER SUMMARY	25
3	CHAPTER 3: PRIOR STUDIES OF TRANSCRIPT REVISION TASKS.....	26
3.1	FINDINGS FROM PRIOR STUDIES	26
3.1.1	<i>Lynch (2001, 2007)</i>	<i>26</i>

3.1.2	<i>Mennim (2003, 2007, 2012)</i>	32
3.1.3	<i>Stillwell et al. (2010)</i>	35
3.2	DISCUSSION AND SUMMARY OF PRIOR OF RESEARCH FINDINGS.....	38
3.2.1	<i>Conclusion and expanded research questions</i>	40
4	CHAPTER 4: METHODS OF DATA COLLECTION	42
4.1	ACTION AND CASE STUDY RESEARCH	42
4.2	RESEARCH SITE AND PARTICIPANTS	43
4.2.1	<i>The course (Interactive English: Speaking and listening)</i>	43
4.2.2	<i>The participants</i>	44
4.3	DATA COLLECTION SCHEDULE AND DATA SOURCES	45
4.4	DATA COLLECTION AND CLASSROOM PROCEDURES	47
4.4.1	<i>Debate Cycle, Lesson 1: Initial delivery of position speech</i>	48
4.4.2	<i>Debate Cycle, Lesson 2: Transcript Revision</i>	52
4.5	TEACHER-STUDENT INTERVIEWS	56
4.6	TRANSCRIPTION PRACTICE AND PILOTING OF PROCEDURES	57
4.6.1	<i>Transcription practice sessions</i>	57
4.6.2	<i>Piloting of Transcript Revision Task</i>	59
5	CHAPTER 5: DATA ANALYSIS METHODS	62
5.1	FIDELITY OF STUDENT SELF-TRANSCRIPTION.....	62
5.2	IDENTIFYING LRES IN PAIR INTERACTION	64
5.2.1	<i>Matching LREs to revisions</i>	68
5.3	ANALYSIS AND CODING OF LRES FROM INITIAL REVISION PAIR-TALK	69
5.3.1	<i>LRE foci</i>	69
5.3.2	<i>LRE outcomes</i>	71
5.3.3	<i>Quality of metatalk</i>	74
5.3.4	<i>Coding for LRE initiation and resolution</i>	76
5.3.5	<i>Inter- and intra-coding of pair-talk data (initial revision)</i>	78

5.3.6	<i>Final count of LREs (initial-revision)</i>	81
5.4	ANALYSIS AND CODING OF LREs FROM FINAL-REVISION PAIR-TALK	82
5.4.1	<i>Identification of LREs</i>	83
5.4.2	<i>Coding of LREs</i>	84
5.4.3	<i>Inter- and Intra-coding of pair-talk data (final revision)</i>	89
5.4.4	<i>Final tally of LREs (final revision)</i>	89
5.5	ANALYSIS FOR RETENTION OF REVISIONS.....	90
5.5.1	<i>Counting of revisions</i>	90
5.5.2	<i>Evidence of retention</i>	90
5.6	COLLECTION AND ANALYSIS OF INTERVIEW DATA	95
6	CHAPTER 6: ENGAGEMENT WITH LANGUAGE DURING TRANSCRIPT REVISION	97
6.1	RESULTS FOR RESEARCH QUESTION 1A: WHOLE CLASS TRENDS DURING INITIAL-REVISION	97
6.1.1	<i>Number of revision-related LREs</i>	97
6.1.2	<i>Foci of LREs</i>	99
6.1.3	<i>LRE outcome</i>	100
6.1.4	<i>Quality of metatalk and LRE outcomes</i>	101
6.1.5	<i>Joint participation in LRE resolution</i>	103
6.2	RESULTS FOR RESEARCH QUESTION 1A: PAIR-BY-PAIR ANALYSIS	104
6.2.1	<i>Pair 1: Naoto and Ken</i>	105
6.2.2	<i>Pair 2: Chika and Momo</i>	110
6.2.3	<i>Pair 3: Naho and Asami</i>	115
6.2.4	<i>Pair 4: Aki and Yuta</i>	117
6.3	RESULTS FOR RESEARCH QUESTION 1B: EWL DURING FINAL-REVISION	120
6.3.1	<i>Pair 1: Naoto and Ken and Pair: 2 Chika and Momo</i>	123
6.3.2	<i>Pair 3: Naho and Asami</i>	131
6.3.3	<i>Pair 4: Aki and Yuta</i>	134
6.4	SUMMARY OF RESULTS	137
6.4.1	<i>Summary of noticing during initial pair revision of transcripts</i>	137

6.4.2	<i>Summary of learner EWL during initial pair revision of transcripts</i>	138
6.4.3	<i>Summary of differences in EWL between pairs</i>	140
6.4.4	<i>Summary of Engagement with teacher-revisions</i>	141
7	CHAPTER 7: ENGAGEMENT WITH LANGUAGE AND RETENTION OF REVISIONS.....	144
7.1	DATA SORTING FOR ANALYSIS.....	144
7.1.1	<i>Sorting of data for analysis of effect of quality of metatalk on retention of revisions</i>	144
7.1.2	<i>Sorting data for analysis of effect of participation in episodes of limited metatalk on retention of revisions</i>	146
7.2	RESULTS	147
7.2.1	<i>Effect of quality of metatalk on retention of revisions.....</i>	147
7.2.2	<i>Participation in limited metatalk and retention of revisions.....</i>	149
7.3	SUMMARY OF RESULTS	150
8	CHAPTER 8: LEARNER PERSPECTIVES ON TRANSCRIPT REVISION TASKS.....	153
8.1	STUDENT PERSPECTIVES ON SELF-TRANSCRIPTION.....	153
8.1.1	<i>Negative aspects of self-transcription</i>	153
8.1.2	<i>Benefits of self-transcription</i>	154
8.2	PERCEPTIONS OF PEER-EDITING OF TRANSCRIPTS	158
8.2.1	<i>Student preferences for future editing with or without a partner</i>	158
8.2.2	<i>Reasons for rejecting the NO PARTNER PLAN and retaining peer-editing.....</i>	159
8.2.3	<i>Reasons for recommending self-editing of transcripts.....</i>	163
8.3	USING THE TARGET-LANGUAGE WHEN EDITING TRANSCRIPTS.....	168
8.4	IMPORTANCE OF REDELIVERING SPEECHES.....	170
8.5	FINAL STUDENT RECOMMENDATIONS.....	171
9	CHAPTER 9 DISCUSSION AND CONCLUSION	174
9.1	EWL DURING TRANSCRIPT REVISION: IMPEDIMENTS AND FACILITATORS.....	175
9.1.1	<i>EWL during initial-revision of transcripts</i>	175

9.1.2	<i>EWL during final-revision of transcripts</i>	<i>180</i>
9.2	EWL AND RETENTION OF REVISIONS	183
9.3	LIMITATIONS OF STUDY AND AVENUES OF FUTURE RESEARCH	184
9.4	PEDAGOGICAL SUGGESTIONS REGARDING USE OF TRTs	187
References		194

APPENDICES

Appendix A: Consent to Participate in Research.....	182
Appendix B: Student profile questionnaire.....	184
Appendix C: Recommended content-organization of debate speeches	185
Appendix D: Example of unrevised student-generated transcript.....	186
Appendix E: End-of-course interview guide.....	188
Appendix F: Transcript conventions for student interaction data.....	192
Appendix G: Full and completed LRE analysis table	193

List of Tables

Table 4.1 Participants.....	45
Table 4.2 Brainstorming teams for position speeches.....	48
Table 4.3 Lesson 1: Initial delivery of position speeches	52
Table 4.4 Revision pairs.....	53
Table 4.5 Lesson 2: Transcript revision	54
Table 4.6 Lesson 3: Redelivery of speeches and debate.....	56
Table 6.1 Number of revision-related LREs across TRTs	98
Table 6.2 Summary of foci of LREs (initial transcript revision)	99
Table 6.3 Foci of Form-LREs and Lexical-LREs (initial transcript revision)	100
Table 6.4 LRE outcomes by focus (initial transcript revision)	101
Table 6.5 Quality of metatalk and LRE outcomes (initial transcript revision)	102
Table 6.6 Quality of metatalk and LRE focus during initial transcript revision	103
Table 6.7 Individual vs. joint initiation/resolution of LREs.....	104
Table 6.8 Pair 1: Individual vs. joint initiation/resolution of LREs	105
Table 6.9 Pair 1: Quality of LRE metatalk across TRTs	110
Table 6.10 Pair 2: Initiation of LREs across three TRTs	111
Table 6.11 Pair 2: Individual vs. joint initiation/resolution of LREs	112
Table 6.12 Pair 2: Quality of LRE metatalk across TRTs	114
Table 6.13 Pair 3: Individual vs. joint initiation/resolution of LREs	116
Table 6.14 Pair 3: Quality of LRE metatalk across TRTs	117
Table 6.15 Pair4: Co-resolution across TRTs.....	118
Table 6.16 Pair 4: Quality of LRE metatalk across TRTs	118
Table 6.17 Noticing of inaccurate language across TRTs	121
Table 6.18 Pairs 1 and 2: EWL with Type 2 teacher-revisions	124
Table 6.19 Pair 1: Metatalk and uptake of Type 3 teacher-revision.....	127
Table 6.20 Pair 2: Metatalk and uptake of Type 3 teacher-revision.....	128
Table 6.21 Pair 3: Metatalk and Type 2 teacher-revisions.....	132
Table 6.22 Pair 3: Metatalk and uptake of Type 3 teacher-revisions	134
Table 6.23 Pair 4: Metatalk and Type 2 teacher-revision	134

Table 6.24 Pair 4: Metatalk and uptake of Type 3 teacher-revision	135
Table 6.25 Comparison of pair EWL during initial revision of transcripts.....	140
Table 7.1 Quality of metatalk and retention of single round revisions	148
Table 7.2 Participation in episodes of limited metatalk and retention of revisions	150
Table 8.1 Positive aspects of self-transcribing own oral performance	156
Table 8.2 Editing pairs and individual student editing preference	159
Table 8.3 Reasons for rejecting <i>NO PARTNER PLAN</i> and retaining peer-editing	160
Table 8.4. Perceived benefits of redelivering speeches.....	171
Table 9.1 Cognitive EWL during initial revision of transcripts	176
Table 9.2 Patterns of participation and during initial pair-revision	177
Table 9.3 Comparison of pair EWL during initial-revision of transcripts	179
Table 9.4 Comparison of pair EWL during initial-revision of transcripts	181

List of Figures

Figure 4.1 Outline of lessons and tasks that comprised a 'Debate Cycle'	46
Figure 4.2 Main study data collection schedule	47
Figure 4.3 Pre-formatted 'debate house' for position speech.....	50
Figure 5.1 Example of LRE matched to revision.....	66
Figure 5.2 Examples of discontinuous and embedded LREs	67
Figure 5.3 Example deletion of dysfluencies.....	68
Figure 5.4 Example LRE identification and analysis table	78
Figure 5.5 LRE identification and analysis tables for independent coding	80
Figure 5.6 Matched sentences: Initial, teacher, and final revisions	83
Figure 5.7 LRE outcome coding for final pair-revision	85
Figure 5.8 LRE coding for pair-talk during final transcript revision.....	87
Figure 5.9 Examples of retained revisions	92
Figure 5.10 Examples of revisions not retained	94
Figure 6.1 Example of learners declining to adopt teacher-revision (Equivalency)	128
Figure 6.2 Example of learners declining to adopt teacher-revision (Difference).....	129
Figure 6.3 Example of learner adoption of teacher-revision (Novelty)	130
Figure 6.4 Pair 4 discussing teacher non-revision of style	135

1 INTRODUCTION

1.1 Supplementing communicative tasks with focus on form activities

This study has its roots in my having been assigned by my university in Japan to teach a 1st year undergraduate course entitled *Interactive English: Speaking & Listening*. I was free to design the course as I saw fit, with the course aims as specified in a memo from my institution being merely:

To effectively link what is taught in high schools to students' productive skills. To improve productive skills in English, we request that teachers adopt various types of productive activities in class such as discussion, speech, and debate.

I elected to center the syllabus on having students debating in pairs with one student assigned to argue for a given resolution and the other against. These pair debates commenced with students delivering speeches: one student delivering a speech in support of the resolution; followed by the other student delivering a speech in opposition. The student pair then proceeded to refuting the points contained in their opponent's speech and defending their own.

In prior iterations of the course, the debate tasks were found to be successful insofar as the students were able to deliver well-reasoned position speeches followed by lively debate. The English produced, however, was often inaccurate. 'Accuracy' here refers to "how well the target-language is produced in relation to the rule-system of the target language" (Skehan, 1996, p.23) and is restricted to the morphosyntactic and lexical accuracy of student production. In light of the inaccuracy of much of the students' oral output, I came to the conclusion that had been doing them a disservice by employing a syllabus comprised almost exclusively of communicative tasks as such syllabi are unlikely to adequately assist learners towards

improving the linguistic accuracy of their production. As Skehan *et al.* (2012, p. 171) note:

The idea that [communicative] tasks in themselves contain all that is needed for sustained second language development has been discredited, and it is now recognized that within tasks, there needs to be some degree of focus-on-form.

However, I felt it unrealistic—even unfair—to expect learners to focus their attention towards accuracy of form during debates when conveyance and comprehension of meaning is of primary importance. Studies such as VanPatten (1990) and Wong (2001) have shown that learners find it difficult to attend to meaning and language form at the same time. Furthermore, although Long (e.g., 1996, 2006) posits that breakdowns in communication during task performance will lead learners to focus on form, studies have found that such breakdowns rarely occur in the classroom (e.g., Ellis, Basturkmen, & Loewen, 2001; Foster, 1998). Additionally, in the absence of breakdowns in communication, learners may simply not feel it socially appropriate to point out errors they notice in their interlocutor's production. Typically, when students perform communicative tasks, "meaning is focused on and error is ignored in an attempt to create an effective social interaction" (Swain, 2000, p.107). Finally, I felt providing learners with teacher corrective-feedback during task performance undesirable, as it would entail my interrupting the flow of learners' debates.

1.2 Transcript Revision Tasks

To draw learner attention towards the accuracy of their language production, this study trialled a novel activity suggested by six studies: Lynch (2001, 2007), Mennim (2003, 2007, 2012), and Stillwell *et al.* (2010). These studies employed what I will label 'Transcript Revision Tasks' (TRTs) which had learners:

1. Self-transcribe, verbatim, audio recordings of their oral performance of a communicative task.
2. Work with peers to revise any language errors noticed in the transcripts.
3. Compare these student-generated revisions to those made by the teacher to the transcripts.
4. Re-perform the original oral task.

In this study, students' deliveries of their individual debate position speeches were recorded. As homework, each student was asked to self-transcribe their speeches and submit these transcripts to the teacher. In the subsequent lesson, students worked in pairs and attempted to identify and revise language errors contained in the transcripts. Pairs then compared their revisions to those the teacher had made to the same transcripts. In the lesson after that, these position speeches were redelivered by the students. (These procedures are specified in greater detail in Chapter 4).

1.3 The purpose of TRTs

The primary purpose of TRTs is to provide learners with opportunities to focus their attention towards the accuracy of the language forms they are producing. Providing learners with these opportunities is intended to promote 'noticing', 'collaborative dialogue', and 'engagement with language'. These concepts are briefly introduced below.

By directing TRTs their focal attention towards language form, TRTs are intended to promote what Schmidt (1990, 1995, 2001) calls 'noticing.' Specifically, TRTs are intended to help learners 'notice the gap' (Schmidt and Frota, 1986) between their own L2 output and target language norms. TRTs do this by providing learners the opportunity to work with peers to identify and repair errors

autonomously from the teacher, and then providing learners with subsequent opportunity to receive feedback from the teacher regarding the accuracy of the language they produce. This process of noticing and repairing errors is meant to promote meta-talk about language between learners, or ‘linguaging’ (Swain, 2006) in the form of ‘collaborative dialogue’ where “learners work together to solve linguistic problems and/or co-construct language or knowledge about language” (Swain, Brooks, and Tocalli-Beller, 2002, p. 172).

Svalberg (2009; 2012) proposes the construct of ‘engagement with language’ (EWL) as a model framework for exploring how learners achieve awareness of form. As “the learning of a language or languages, normally relies on a combination of cognitive, affective, and social factors” (Svalberg, 2009, p. 249), EWL is a threefold construct which encompasses cognitive, social, and affective engagement. The EWL model posits that extent and quality of attention learners pay to form will depend on of how cognitively, affectively, and socially engaged learners are during tasks where language form is the intended object of study.

1.4 Limitations of prior investigations into classroom use of TRTs

Previous studies reporting on the class-room use of TRTs—especially those conducted at Japanese universities with learners comparable to my own (Mennim, 2003, 2007, 2012; Stillwell *et al.*, 2010)—were of great value insofar as they established TRTs as a feasible means of focusing my learners’ attention to form. Previous TRT studies, however, have primarily examined the products of transcript revision; for example, which language features learners revised and the how accurate these revisions were; but limitations in the studies’ classroom data collection procedures (discussed in Chapter 3) often make these results difficult to interpret. The main limitation of previous TRT studies, however, is that they provide little discussion about the process of noticing and collaborative dialogue that went on between learners, or whether learners saw value in self-transcribing and

collaborating with others when making revisions. In short, these studies tell us little about how effective the TRTs were at meeting what this study argues is their primary purpose: to promote learner EWL.

1.5 Purpose of study and research questions

The purpose of this study is to evaluate the efficacy of the TRTs trialled in the debate course in terms of the extent and quality of the learner EWL generated by the tasks. Situated within prior SLA literature and research on noticing, collaborative dialogue, engagement with language, and TRTs, this study investigates: (a) the quality of learner EWL during transcript revision; (b) the effect differences in EWL amongst learners had on retention of revisions; and (c) the learners' perceptions of TRTs as a language learning experience. Three research questions guided the thesis study in the context of an English course at a university in Japan:

1. What was the quality of learner EWL during transcript revision?
2. Did differences in EWL during transcript revision affect retention of revisions?
3. How did the learners perceive Transcript-Revision-Tasks as a language learning opportunity and experience?

1.6 Organization of the thesis

This thesis is organized into eight subsequent chapters. Chapter 2 (the following chapter) presents the conceptual background and discusses the importance for SLA of: (i) noticing of input and output, (ii) peer-peer discussion of L2 form, and (iii) engagement with language. Chapter 3 reviews previous literature on transcript revision tasks. Chapter 4 outlines the research design of the study and describes: (a) the research site and context, (b) the participants, (c) data sources, (d) piloting of data collection procedures, (d) the finalized data collection procedures. Chapter 5

outlines the methods of data analysis. Chapters 6 to 8 report and briefly discuss the findings with respect to the three research questions. Chapter 9 concludes the thesis. In that chapter, the findings are synthesized and discussed in light of relevant prior research. Next, the limitations of the study and suggested avenues for future research are presented. Lastly, the chapter concludes with the pedagogic implications of the findings and suggests ways in which TRTs could be better designed and implemented to potentially improve the degree of learner EWL.

2 CHAPTER 2: CONCEPTUAL BACKGROUND

Transcript revision tasks are intended to generate learner attention to form. Specifically, the tasks intend to promote *noticing*, *peer-peer discussion of form*, and *engagement with language*. These three concepts-cum-processes are the focus of this chapter.

2.1 Noticing

Schmidt's (1990, 2001) 'Noticing Hypothesis' addresses the debate regarding the type of attention necessary for language input to become intake. Input is defined here as, "the potentially processable language data which are made available by chance or by design, to the language learner" (Sharwood Smith, 1993, p. 167); and intake as, "the subset of input which becomes accessible to whatever the key process are of learning" (Reinders, 2010, p. 62).

Unpacking the concept of 'noticing' requires brief discussion of its underlying concepts of attention and awareness. While there is disagreement as to the precise meaning of attention (Segalowitz and Lightbown, 1999), the distinction made between perceptual and focal attention is widely recognized (Dörnyeni, 2009; Robinson, 2015). Perceptual attention is paid automatically, involuntarily, and unconsciously to stimuli from the environment (Robinson, 2015); whereas, focal attention is subject to voluntary control and refers to "the ability to concentrate on certain stimuli while ignoring others" (Dörnyeni, 2009, p. 133). 'Awareness' is another concept with no universally agreed upon definition, but most definitions include the idea that awareness entails one's knowledge and subjective experience of having detected a stimulus (Al-Hejin, 2004). Focal attention is generally agreed to be a precondition of awareness (Robinson, 2015), whereas the mere perception of stimuli by the senses (i.e., perceptual attention) can be dissociated from awareness (Tomlin and Villa, 1994).

2.1.1 The Noticing Hypothesis

In SLA, there is dispute whether perceptual attention is sufficient for input to become intake or whether focal attention-*cum*- awareness is required. The Noticing Hypothesis claimed that focal attention was required for input to become intake. According to Schmidt (1990), input which is merely perceived would persist only fleetingly in immediate short-term memory and be almost immediately lost. Rather, the hypothesis claimed that conversion of input into intake required 'noticing', which is the learner's subjective awareness that they had detected a specific L2 feature in the input. Input that was noticed could become encoded in short-term memory and thus available for further cognitive processing. Schmidt was careful to distinguish between awareness at the level of 'noticing' from 'understanding' which is a higher-level of awareness of an abstract L2 rule or principle of which the input feature may be an exemplar. While awareness at the more abstract level of understanding could facilitate learning, the hypothesis posited that, "noticing is the necessary and sufficient condition for converting input to intake" (1990, p. 129).

2.1.2 Noticing and awareness as facilitators of language acquisition

While Schmidt's argument that awareness is required for learning has garnered support (e.g., Ellis, 1994, 1997, 2003; Robinson, 1995, 2003; Skehan, 1998); this support has not been universal (e.g., Tomlin and Villa 1994; J. N. Williams, 2005). Schmidt (2001) moderated his position somewhat regarding the impossibility of learning without awareness to "people learn about the things they attend to and do not learn much from the things they do not attend to" (p. 30). As Schmidt rightfully noted, for SLA pedagogy, the more important question was not whether any learning at all could take place without attention and awareness, but rather "whether more attention resulted in more learning" (2001, p.28).

A series of studies by Leow (1997, 2000, 2001), Rosa and Leow (2004), and Rosa and O'Neill (1999) have demonstrated that different degrees of awareness

result in different degrees of learning. These studies shared the following design features. Individual participants performed problem solving puzzles (e.g., crossword puzzles in the case of Leow's studies) targeting aspects of Spanish grammar (e.g., Spanish contrary-to-fact conditionals in Rosa and Leow, 2004). Pretesting was employed to ensure participants had no prior knowledge of the targeted grammatical feature. Think-aloud protocols were used in each study to gauge the level of learners' awareness of targeted forms during task performance. In these studies, learner awareness was categorized into three levels:

- (i) *Noticing* in cases of learner verbal reference to the target structure without any mention of rules.
- (ii) *Understanding* in cases where learners attempted to formulate a target structure rule.
- (iii) *No verbal report* for learners who limited themselves to reading aloud puzzle clues without verbal signal that the target structure or the rule governing it had been registered. (As absence of self-report does not necessarily imply lack of awareness, the studies made no claims regarding the awareness of learners in the *no verbal report* category).

The above studies found that learners who demonstrated awareness of targeted form at the level of 'noticing' experienced a score growth on post-tests significantly larger than learners who did not (i.e., 'no verbal report'). Additionally, learners who demonstrated awareness at the higher level of 'understanding' had significantly higher scores than learners who demonstrated only noticing. Rosa and Leow (2004) found that these effects held in both immediate and delayed post-tests administered 3 weeks later. Leow (1997, 2001) further reported that learner demonstration of noticing correlated with increased use of hypothesis testing and rule formation (i.e., 'understanding'); but absence of noticing correlated with

absence of hypothesizing and rule formation. In other words, it appeared that noticing often triggered cognitive processes that led to understanding and was the precursor of these processes.

2.1.3 The Output Hypothesis and noticing

Swain (1993, 1995) in her 'Output Hypothesis' pointed out that while it was possible for learners to comprehend input without necessarily having to notice form-features or process the input linguistically (for example, learners could use context to guess meaning), producing comprehensible output required syntactical processing. Swain further argued that when producing output learners may notice limitations in their L2 knowledge.

Firstly, learners may 'notice a hole' in their L2 knowledge: "even without implicit or explicit feedback provided from an interlocutor about the learners' output, learners may still, on occasion, notice a gap in their own knowledge when they encounter a problem in trying to produce the L2" (Swain and Lapkin, 1995, p. 373). Noticing such a hole in one's L2 knowledge may then lead the learner to either pay greater attention to relevant input or search their own linguistic knowledge to find information which can help close the hole. Swain argued these processes could then "lead to creation of new L2 knowledge or consolidation of existing knowledge" (1993, p. 159).

Secondly, as a learner's output represents their best hypothesis as to how something should be said or written in the L2, an interlocutor's reaction to the output can indicate to the learner that their hypothesis is erroneous (Swain, 1995). Such interlocutor feedback, whether implicit or explicit, would lead the learner to 'notice the gap' between their output and the TL norms (Schmidt and Frota, 1986). Noticing the gap has been argued to lead learners to restructure their developing L2 towards TL norms (e.g., Long, 2007; Lyster and Ranta, 1997; Swain, 1995, 1998, 2005; Swain & Lapkin, 1998).

Finally, Swain posited that learners' attempts to produce comprehensible and accurate output can trigger a 'meta-linguistic' function "when a learner explicitly reflects on their hypotheses about the L2 using language to do so" (1995, p. 132). In other words, when a learner notices they are unsure of how to correctly express themselves in the L2, they may proceed to experiment with possible ways of using a structure, trial alternative structures, or seek an explanation for why a form is correct or not. As discussed earlier in section 2.1.2, noticing that resulted in deeper processing of language data, such as hypothesis testing or rule formation, was found to result in superior learning gains.

2.2 Peer-peer discussion of language

The Noticing and Output hypotheses are underpinned by cognitivist perspectives on SLA which view acquisition as a matter of an individual, internal, mental phenomenon (Ellis, 2015). However, since putting forward the Output hypothesis, Swain (e.g., Swain, 2000, 2006, 2010) and her co-researchers (e.g., Kowal and Swain, 1994; Lapkin, Swain, and Smith, 2002; Tocalli-Beller and Swain, 2007; Watanabe and Swain, 2007) have taken a sociocultural theory (SCT) orientation towards L2 development to argue the importance of peer-peer meta-talk triggered by noticing. While not all studies of learner meta-talk make reference to SCT (e.g., Ecketh, 2008; J. Williams, 2001), a great number are underpinned by SCT. Therefore, this section briefly introduces SCT before turning to discussion of peer-peer deliberation over language.

2.2.1 Sociocultural theory

Sociocultural theory (SCT) is based upon the work of Soviet psychologist and educator L. S. Vygotsky. The core premise of SCT is that, while biology endows humans with the capacity for higher cognitive functions, "all uniquely human forms of mental activity arise solely as a consequence of the dialectical interconnection between what is provided by nature and what is offered by human culture" (Lantolf,

2006, p. 103). According to Vygotsky, higher cognitive functions first appear on an inter-psychological plane during social interaction between individuals which then subsequently become internalized by the individual on an intra-psychological plane (Vygotsky, 1978 cited in Lantolf, 2000a, 2000b). In SCT, this 'internalization' constitutes an individual's cognitive development (Lantolf, 2005), and is the result of connecting the social and individual planes when knowledge constructed socially becomes transformed into an internal resource the individual can employ independently (Lantolf and Thorne, 2007). Language holds a special place in SCT in that it is both a cultural artifact that mediates mental functions and simultaneously a uniquely human form of mental activity which itself undergoes social/cultural development (Kozulin, 1986, p. xxx).

2.2.2 *The Zone of Proximal Development*

SCT further argues that knowledge can only be constructed and internalized when the assistance a learner receives during social interaction falls within the learner's 'zone of proximal development' (ZDP): which is, "the difference between what an individual can achieve on their own and what the individual can accomplish during problem solving with the support and guidance of a more capable teacher or peer" (Lantolf, 2000a, p. 17). A suggested definition of ZDP for a language learning context is:

...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)

However, the ZPD is not an attribute of the learner *per se*, but rather something discovered and co-constructed through participation in collaborative dialogic activity (Lantolf, 2000b). As Poehner (2008) points out, "successful collaboration in the ZPD

is dependent upon both the quality of mediation and learner reciprocity” (p. 40). If, for example, the mediator simply solves a language problem on the learner’s behalf, the construction of the learner’s ZPD is unlikely to occur (Ohta, 2000). In this sense, Swain, Kinnear, and Steinmann (2011) suggest it may be helpful to view the ZPD as an activity in which both expert and novice participate.

SLA research has demonstrated that when learners work together to resolve language problems, they can concurrently be experts and novices (e.g., Donato, 1994; Ohta, 2000, 2001; Swain and Lapkin, 1998; Watanabe, 2008; Watanabe and Swain, 2007). As Ohta (2001) explains, “No learner is universally more or less capable than a peer, but that each learner presents an array of strengths and weaknesses that may be complementary” (p. 76). By pooling linguistic resources—what Donato (1994) labels “collective scaffolding”—learners who are “individually novices” can become “collectively experts” (Donato, 1994, p. 46).

2.2.3 *Collaborative Dialogue*

Swain (2000) coined the term ‘collaborative dialogue’ to describe instances where learners work together to solve linguistic problems they encounter when producing L2 output:

...collaborative dialogue is problem-solving and, hence, knowledge-building dialogue. When a collaborative effort is being made by participants in an activity, their speaking (or writing) mediates this effort. As each participant speaks, their “saying” becomes “what they said,” providing an object for reflection. Their “saying” is cognitive activity, and “what is said” is an outcome of that activity. Through saying and reflecting on what was said, new knowledge is constructed. (Swain, 2000, p. 113)

Collaborative dialogue entails ‘linguaging’, a term Swain uses to cover the meta-talk process “of making meaning and shaping knowledge and experience through language” (2006, p. 98). Linguaging can be entirely self-directed, for example, when a learner talks to themselves whilst working alone on a complex language task (see, for example, Knouzi et al., 2010; Swain et al., 2009). When produced while working with peers, self-directed linguaging becomes public, and interlocutors may then, for example, ask the learner to further explain or justify their ideas or see a learner’s self-directed questions as signals for assistance (Swain and Lapkin, 2002). In other words, self-directed linguaging may initiate collaborative dialogue.

2.2.4 Language related episodes

In SLA research, collaborative dialogue is usually examined in relation to ‘language-related episodes’ (LREs) which are instances where “students talk about the language they are producing, question their language use, or other- or self-correct” (Swain, 1998, p. 70). The frequency and linguistic focus of LREs appears tied to the overall proficiency of the dyad engaged. For example, Leeson (2004) found that the high-high proficiency dyads in his study produced far more LREs than the low-low and high-low dyads during performance of a dictagloss task. Similarly, Kim and McDonough (2008), Watanabe and Swain (2007), J. Williams (1999, 2001) all reported that the total number of LREs tended to be higher when the overall proficiency level of dyads was greater.

Research also indicates that the majority of learners’ LREs focus on lexis or grammatical form (e.g., Kim and McDonough, 2011; Leeson, 2004; Malmqvist, 2005; Storch, 2008; Swain and Lapkin, 2002; J. Williams, 1999, 2001). These findings are unsurprising given that, “To make their meaning as clear, coherent and precise as possible, learners will debate language form (morphosyntax through to discourse and pragmatics) and lexical choice” (Swain and Lapkin, 2002. p. 286). The studies also examined LREs in terms of whether the LREs were resolved correctly,

incorrectly, or left unresolved. These studies found that, in general, the higher the overall proficiency of a dyad, the more frequently LREs were resolved correctly. The high-high proficiency dyads in Leeson's (2004), for example, were able to correctly resolve 88% of LREs compared to 64% for high-low pairs and 58% for low-low dyads (who abandoned and left 33% of their LREs unresolved).

Linking collaborative dialogue to L2 development

Studies investigating the contribution of collaborative dialoguing to L2 development often employ tailor-made posttests or a process-product research design. Tailor-made posttest items are, therefore, developed *a posteriori* after the collaborative talk has been analysed for the foci of the LREs. However, as Storch (2013) points out, creation of dyad/group specific posttest items is not always feasible within a study's time constraints for data collection (for example, in classroom-based studies with a larger number of students). Therefore, a process-product design is also common in studies of collaborative dialogue. In this approach, researchers set out to determine whether the L2 forms that were the foci of collaborative dialogue are used by learners in subsequent individual performance of similar (isomorphic) tasks or when repeating the same task.

Studies employing tailor-made posttests (e.g., Eckerth, 2008; Kim, 2008; McDonough and Sunitham, 2009; Swain, 1998; Swain and Lapkin, 2002; Tocalli-Beller and Swain, 2007; J. Williams, 2001), and also those utilizing a process-product design (e.g., Brooks and Swain, 2009; Donato, 1994; Storch, 2002, 2008; Storch and Wigglesworth, 2010), have found learners able to convert the L2 lexical and grammatical knowledge they correctly co-constructed into accurate performance on tailor-made posttests or similar/repeated tasks. However, Swain (1998) and J. Williams (2001), for example, also report a close connection between dyads' incorrectly resolved LREs and incorrect answers to items in tailor-made posttests.

Storch (2002, 2008) reports similar instances of incorrect LRE resolutions being transferred to subsequent tasks performed individually.

The above findings strongly indicate that peer-peer languaging results in L2 learning, but also the need for teacher confirmation of LRE resolutions to ensure learning proceeds in a “desired direction” (Storch, 2002). Granted, it is hard to determine if results from tailor-made post-tests or subsequent individual performance of similar/same tasks “represent the incorporation of new linguistic knowledge into the learner’s interlanguage system or if they represent a consolidation of previously existing knowledge” (Loewen, 2005, p. 382); but as Swain and Lapkin (1998) argue, both construction of new knowledge and consolidation of existing L2 knowledge are important aspects of language development. As will be shown below, however, many of the studies cited in this section have additionally found learning outcomes to be dependent on how learners ‘engage with language’.

2.3 Engagement with language

Engaging in collaborative dialogue is a means by which learners can develop Language Awareness (LA), defined here as, “explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching, and language use” (from the Association for Language Awareness <www.languageawareness.org>, cited in Svalberg 2012, p. 376). As noted by Svalberg (2007), “As collectively constructed over the last 15–20 years, LA does not refer to a purely intellectual awareness and is not passive” (p. 302). Rather, Svalberg (2007, 2009, 2012) argues, LA should be viewed as a product and process that arises as a result of ‘Engagement with language’ (EWL).

EWL is a model framework recently proposed by Svalberg (2009, 2012) in which to investigate the cognitive, social, and affective factors that affect (positively or negatively) the attention learners pay towards language. EWL is defined as:

In the context of language learning and use, Engagement with language (Engagement) is a cognitive, and/or affective, and/or social state and a process in which the learner is the agent and language is the object (and sometimes vehicle) (Svalberg, 2009, p. 247).

The EWL framework is intended as a means for researchers and practitioners to construct a picture of what “goes on” during the building of language awareness and “provide a principled way of establishing, measuring, and interpreting the presence/absence, degree, and nature of Engagement” (Svalberg, 2009, p. 256). Individuals engaged with language are “actively constructing their knowledge not only by mental processes but also equally by being socially active and taking initiatives” (ibid, p. 246). For teachers employing tasks intended to lead their learners to engage in collaborative dialogue, the EWL framework can be a means for ascertaining what cognitive, social, or affective factors are facilitating or impeding such dialogue. EWL is intended to provide a framework for a more holistic exploration of how language awareness is constructed and Svalberg concedes that the division of EWL into cognitive, social, affective sub-components is artificial. However, as Svalberg argues, “for the purpose of analysis it seems useful to separate facilitators and impediments of EWL according to the three types of Engagement outcomes they affect” (2009, p. 255).

2.3.1 Cognitive engagement

Svalberg posits that, cognitively speaking, an individual Engaged with language “is alert, pays focused attention and constructs their own knowledge” (2009, p. 246), and that a state of heightened cognitive engagement will manifest in a process of focused reflection and problem solving (p. 255). However, as shown below, task

demands may largely determine the extent to which problem solving and extensive reflection arises.

Storch (2008) investigated the effect the quality of learner cognitive engagement with linguistic choices had on LRE resolution. 11 pairs of advanced ESL learners completed a text reconstruction task (i.e., attempted to produce a meaningful and grammatically accurate text by inserting missing function words and changing word forms). LREs from the recorded pair-talk data were coded in terms of their grammatical/lexical foci and whether the LREs evidenced 'elaborate' or 'limited' cognitive engagement. Elaborate engagement was operationalized as instances of joint LRE resolution involving deliberation and discussion of language items. Limited engagement was operationalized as instances where LREs were resolved without deliberation or discussion by a single learner with the other learner merely repeating, acknowledging, or not explicitly responding to the suggestion. The study found that the majority of LREs were resolved correctly regardless of the nature of the cognitive engagement: 80.3% of LREs showing elaborate engagement and 79.44% showing limited engagement were resolved correctly. Supplying the correct text amendments often appeared unproblematic; for example, only 24% (16/67) of LREs involving choice of prepositions showed elaborate engagement, which Storch observed, "were fairly easy for these advanced learners" (p. 108). Not all language 'problems' will require extensive cognitive engagement and discussion to resolve correctly.

Storch and Wigglesworth (2010) compared how learners discussed direct teacher corrective-feedback (reformulations) versus indirect feedback (editing symbols) directed towards language errors contained in the learners' pair compositions. Similar to Storch (2008), the study drew a distinction between LREs that showed 'extensive' engagement (e.g., deliberation or meta-talk about the feedback) and those that showed limited engagement. Editing symbols were found

to have elicited far more incidences of extensive cognitive engagement than direct teacher feedback (reformulation) which was found to usually generate limited engagement where one member of the pair simply read the feedback and the other repeated/acknowledged it. The researchers suggest the differences in cognitive engagement can be accounted by the fact that indirect feedback still demands learners formulate corrections themselves; whereas, as one pair of learners was recorded putting it, reformulations “give away the answers” and leave learners nothing to do other than “just memorize” the feedback. In short, direct feedback appeared to have left learners less room or reason to construct their own knowledge.

Storch (2008) and Storch and Wigglesworth (2010), in addition to illustrating how task demands affected the quality cognitive engagement, also investigated the effect the quality of cognitive engagement with linguistic choices had on subsequent learning outcomes. Both studies employed a process-product research design with instances of language amendments or revisions (whether correct or incorrect) made to texts by pairs subsequently amended/revised the same way on a similar or same task performed individually taken as evidence of language learning. The overall finding from both studies was that elaborate/extensive engagement resulted in more instances of learning than did limited engagement. While both studies caution that coding for engagement is highly inferential and the amount a learner verbalizes may not necessarily reflect the depth of his/her cognitive processing, their findings indicate that more extensive meta-talk leads to superior learning gains.

Storch’s (2008) study, however, also suggests that the connection between quality of meta-talk and learning is more nuanced. Storch found in slightly over half the cases where LRE resolution involved elaborate engagement that both learners showed evidence of learning. In cases of limited engagement, however, the study

found the learning benefits most often accrued to the learner who suggested the language amendment rather than to the other who merely accepted it. Although, as previously noted, not all LREs required deliberation to resolve, these findings indicated that the learner who appeared more alert and initiated and resolved an LRE was the one most often reaping the learning benefit.

2.3.2 Social Engagement

Collaboration in a learning environment involves building and fostering social relationships (Donato, 1994). Social engagement refers to the quality of students' participation in the classroom which is "essentially linked to interaction and to learners' initiation and maintenance (or not) of it" (Svalberg, 2009, p. 252). Svalberg posits that socially, a learner fully engaged with language is interactive and initiating (2009, p. 247).

The importance of learners being initiating and interactive is highlighted by studies that have employed Storch's (2002) framework for identifying 'patterns of dyadic interaction'. Storch's framework looks at two key aspects of interaction: equality and mutuality. 'Equality' refers "to the degree of control or authority over the task" with higher degrees of equality evidenced by "interactions where both participants take directions from each other" (Storch, 2002, p. 127). 'Mutuality' refers to the amount learners engage with each other's contributions and "interactions that are rich in reciprocal feedback and a sharing of ideas" characterize high mutuality (ibid.).

Using these indices of equality and mutuality, Storch identified four patterns of dyadic interaction:

- I. The collaborative pattern evidenced relatively high degrees of both equality and mutuality and emerged when both learners in a pair were initiating and interactive. Learners were relatively equal in terms of initiation of discussions

with both learners willing to engage with each other's ideas, to provide feedback, explanations and suggest alternative suggestions.

- II. The expert/novice pattern evidenced relatively low equality with one pair member assuming or being afforded the role of expert who led the task by initiating most discussions. However, this pattern also showed relatively high mutuality with both participants being interactive. The expert actively encouraged the 'novice' learner to take part in discussion and supported the novice's learning by providing explanations or other scaffolded assistance; novices responded to experts' invitations and repeated/confirmed the expert's suggestions/explanations.
- III. The dominant/passive pattern evidenced low levels of both equality and mutuality. The dominant member of such dyads appropriated the task; initiated most suggestions; and made little attempt to involve the other learner or seek their contributions. Little negotiation or discussion ensued because the passive participant made few challenges or contributions (usually limited to mere echoic repetitions of the dominant member's suggestions).
- IV. The dominant/dominant pattern evidenced high equality with both learners initiating meta-talk and making suggestions, but low mutuality as learners appeared unwilling to engage with each other's suggestions. Discourse was marked by disagreement and learners tended to disregard each other's suggestions.

Storch (2002) found that learners whose interactions evidenced a collaborative or expert/novice pattern showed more transfer of knowledge to subsequent individually performed tasks than did learners whose interactions were dominant/passive or dominant/dominant. Opportunities to engage in solving

language problems were often missed by the dominant/dominant pair and arose only for the dominant learner—whose meta-talk was self-directed—in the dominant/passive dyad. That is, when one learner adopted a dominant stance, the other lost opportunities to engage in the kind of language learning that occurred amongst collaborative or expert/novice dyads. However, Watanabe and Swain (2007) found a clear learning advantage for both learners only in pairs whose interactions evidenced the collaborative pattern. In expert/novice pairs, only the experts, but not the novice partners, had performed better on post-tests.

As Storch (2002) notes, while a collaborative pattern of interaction best fosters learning; this pattern does not necessarily arise just because learners are asked to work together. Svalberg (2009) posits that among learners of unequal L2 proficiency, this ‘power differential’ may impede social EWL. However, studies examining how learner interactions change when they work with a higher- versus a lower proficiency partner have produced mixed findings. While Watanabe and Swain (2007) and Storch and Aldosari (2013) found that in expert/novice or dominant/passive dyads it was the relatively lower proficiency learner who was the novice or passive participant, these studies also observed mixed-proficiency pairs adopting collaborative patterns of interaction. Kim and McDonough (2008) found intermediate learners adopting collaborative stances when working with another intermediate interlocutor who took passive or novice stances when paired with a more advanced speaker, but also other intermediate learners adopting dominant roles with other intermediate learners who took collaborative stances when working with advanced partners. Proficiency differences in themselves, therefore, do not appear to necessarily determine the quality of learners’ social engagement.

2.3.3 Affective Engagement

Svalberg (2009) posits that, affectively, individuals engaged with language will possess a “positive orientation towards the language, the interlocutor, and/or what

they represent” and whose “willingness to interact with the language and/or interlocutor is maintained/heightened” throughout the EWL process (p. 255). She further posits that affectively the engaged individual those engaged will have a purposeful and autonomous disposition and able to take charge of their own learning (p. 246).

That a learner’s orientations can either impede or facilitate EWL is clearly demonstrated in Sato (2017) who investigated how learners’ predispositions towards task and interlocutors impacted languaging between peers. The study involved 10 Chilean high school EFL learners who each had been interviewed about their attitudes towards L2 group work (e.g., Do you think you and your classmates can help each other to learn English?); peer corrective feedback (e.g., Do you feel comfortable correcting your classmate’s mistakes?); and interacting in the L2 (e.g., Do you enjoy talking to your classmates in English?). These learners then performed a series of collaborative language-focused tasks in two groups of five. In one of the groups, four out of the five learners had expressed in the prior interviews that they felt group-work fun to be fun and beneficial for L2 learning; the remaining student, however, felt group work in the L2 to be socially awkward. Analysis of this group’s interaction found that the four learners who had indicated a positive orientation towards group-work evidenced fluid turn taking, collective scaffolding, and engagement with each other’s suggestions and feedback. The remaining student, in contrast, did not participate at all in discussions about language (although it is not specified what, if any, efforts the others had made to include her). The other group of 5 students, however, was comprised of students had expressed a universal skepticism regarding the learning benefits of L2 group-work. Compared to the first group, these learners engaged in discussions of language far less frequently and displayed marked disengagement with each other’s ideas and feedback.

Sato's (2017) study demonstrates that a learner's predispositions affect willingness to engage with language (what he called their 'interaction mindset') and Svalberg (2009) posits a purposeful predisposition towards language study as a facilitator of EWL. Other affective facilitators and impediments of EWL, however, may only emerge during task performance. Storch (2004), in a follow up to her (2002) study on the patterns of interaction, found that peer collaboration required that learners have a shared or compatible understanding of the task's purpose. Interview data revealed the learners in the collaborative and expert/novice pairs all perceived the purpose of the assigned tasks as having been to contribute to task completion to the best of their ability and share resources and help each other to do so. In contrast, the overriding goal for both learners in the dominant/dominant dyad was to display their own L2 knowledge which led to interaction that was competitive rather than collaborative. As for the dominant/passive pair, the dominant participant's purpose was to complete the task as efficiently as possible, and saw appropriating the task as the best means of doing so. The passive interlocutor, on the other hand, had no clear idea about what her role should be in the tasks.

A lack of trust between learners has been shown to impede EWL. Watanabe and Swain (2007) report the case of one learner who interacted collaboratively when discussing language with a more proficient partner, but who dominated interaction when paired with a less proficient learner. Post-task interviews, reported in Watanabe (2008), revealed that this learner's unwillingness to engage with his partner lower proficiency was attributable to his having little trust the partner's ability to play a legitimate role during their interaction. Conversely, learners are unlikely to engage if they feel their input is not valued by their task partners.

Finally, tasks which do not allow learners to exercise their autonomy are less likely to promote ongoing engagement with language: 'autonomy' defined here as,

“the ability to take charge of one’s own learning” (Holec, 1981, cited in Little, 2007, p. 15). Learners who dominate discourse deprive others of the opportunity to take charge of their own learning. For example, some intermediate-level learners in Kim and McDonough’s (2008) study had negative perceptions about working with partners of advanced proficiency because they felt the advanced learners tended to dominate conversations and deprived them of opportunities to participate in task. In contrast, during collaborative interaction, where both learners actively talk about language problems with each other, both learners can exercise their agency as an individual “who perceives, analyses, rejects or accepts solutions offered, makes decisions and so on” (Swain, 2006, p. 101).

2.4 Chapter Summary

Whether learners are able to acquire L2 features without consciously noticing them is still subject to debate, it is generally accepted that noticing is facilitative of acquisition (Ellis, 2015). Noticing can lead to meta-linguistic reflection resulting in deeper processing of language data. Sociocultural SLA researchers posit that such reflection in the form of peer-peer discussion of language can lead to co-construction of new knowledge or refinement of learners’ knowledge of L2 forms. However, more extensive and collaborative discussions of language have been found to result in superior learning gains than those which are perfunctory or one-sided. The quality of learner discussion will depend on—and be a reflection of—the extent of learners’ engagement with language (EWL). Therefore, the efficacy of tasks, especially those such as Transcript-revision-tasks specifically intended to stimulate noticing and languaging, cannot be evaluated without examination of the cognitive, social, and affective factors that facilitated or impeded learner engagement with the task..

3 CHAPTER 3: PRIOR STUDIES OF TRANSCRIPT REVISION TASKS

To date, only a handful of studies have investigated use of transcript revision tasks as a classroom activity. As these studies were my inspiration to the trial incorporation of TRTs into my own course, this chapter begins with an overview of the previous research findings. The second part of the chapter summarizes these findings and discusses their limitations. The chapter concludes with an expanded set of research questions.

3.1 Findings from prior studies

3.1.1 Lynch (2001, 2007)

Lynch (2001), to the best of my knowledge, was the first study to examine TRTs as an L2 learning activity. The study was conducted in the United Kingdom with eight learners enrolled in a university English for Academic Purposes (EAP) oral communication course. The learners were from six different countries, and their average proficiency was approximately 5.5 IELTS. Classroom activities included role-plays performed in pairs in front of the class and video-recorded. This study invited student to participate in an after-class activity. Each pair was given a recording of their role-play performance, and asked to transcribe verbatim a self-selected extract of between 90 and 120 seconds of the recording of their in-class role-play which was called Transcript 1. After that, the researcher asked pairs to make changes to Transcript 1 until they were satisfied with the language to produce Transcript 2. As pairs were transcribing and revising, they were videotaped. At the end of the 45 minute session, the researcher took away the revised transcripts and overnight reformulated any parts that were linguistically incorrect. This teacher-reformulated version was Transcript 3. The next day, pairs were given Transcripts 2 and 3 to compare and discuss with the teacher. The study investigated: (1) the students' attitudes to the transcribing/revising task; (2) whether changes were initiated equally between student who had produced the error (self-correction) and their

revision partner (other-correction); (3) the types of changes and revisions students made; and (4) the proportion of errors produced that students were able to identify.

The study reported that the students appeared to find the task useful and interesting. Lynch reported that pairs “showed no obvious signs of boredom or frustration” when transcribing the role-play extracts, and that at the end of the study “the students asked me to the same thing again with their next set of classroom recordings” (p. 130). The researcher further reported that students cooperated well with each other as they worked to produce revised transcripts. To support this finding, the study presented a single extract from one pair’s interaction during transcript revision. The extract consisted of 18 turns where the pair deliberated over making three changes to the original role-play utterance “*I had to do a lot of work to consult dictionaries*” which they revised into “*I had a lot of work to do consulting the dictionaries*”(changes underlined). The extract clearly showed signs of deliberation, suggestion/counter-suggestion, and collaborative revision.

However, Lynch’s claim that the example extract was “typical of the learner conversations I recorded during the study” (p. 128) was undermined by the study’s findings regarding whether changes were initiated equally between students in pairs. Lynch found that in three of the pairs, changes were initiated almost equally between the students who had originally produced the error and the revision partner; however, in the fourth pair, one student assumed a dominant role and initiated 73.33% (11/15) of revisions. Lynch did not supply any further examination of the fourth pair’s interaction; but, as discussed in Chapter 2, prior studies of dyadic interaction have demonstrated that language problems are seldom resolved cooperatively in pairs where one member was dominant.

Lynch demonstrated that the TRT could offer a productive route to noticing. The four pairs noticed on average 28 points for change in the transcript, which is about one change for every four seconds of speech, and reported that “the time and

trouble they take over details is striking" (2001 p.128). The study used five categories to catalogue learners' revisions:

1. Grammatical corrections, which comprised 35.71% (40/112) of revisions.
2. Lexical correction, which comprised 7.14% (8/112) of revisions.
3. Editing, which comprised 19.64% (22/112) of revisions.
4. Reformulation which comprised 19.64% (22/112) of revisions.
5. Mixed, which comprised 17.85% (20/112) of learners' revisions.

The study found that grammatical corrections were predominantly to verb tenses and choice of articles. 'Editing' referred to "removal of redundancies, literal repetitions, and the sort of dysfluencies and false starts that are investible in natural speech"; and 'reformulation' to "changes made to achieve more precise expression, and insertions of information in order to clarify meanings" (p.129). Note, however, that no examples from these categories were provided nor was the category of 'mixed' defined by the study.

Most of the learners' revisions were reported as being "for the better". Of the 112 revisions made by the four pairs, 72.32% were described as being "for the better"; 17.81% were changes to something correct into an equally correct alternative; and only 9.82% were changes "for the worse" (p. 128). Lynch noted that half of the changes for the worse were made by Pair 4 which, perhaps tellingly, was also the pair in which one member dominated the initiation of revisions. In all, Lynch noted that all pairs noticed and modified about 60% of the points in need of correction, with the remaining 40% being corrected by the teacher. Students were found to have made few vocabulary corrections: only 7.14% of student revisions were to lexical items compared to 32.55% (28/86) of the revisions made by the teacher. Lynch concluded the study by advocating TRTs as a pedagogic activity, noting that the transcription/revision activity yielded several opportunities for

noticing: the process of transcription, reflective self-correction and collaboration with a peer, and discussion of teacher feedback.

Lynch (2007) was a follow-up study conducted in the same university EAP setting, and the material for the study was also generated from pair role-plays. Unlike the previous study, however, TRTs were conducted as in-class rather than as after-class activities. This study involved 16 learners from 10 different countries divided between two intact classes (8 students per class) held once a week for 90 minutes. The purpose of the study was to trial two different TRT procedures: the first involved pairs transcribing their own role-play performances and then revising language errors; in the second procedure, the teacher transcribed the portions of the recordings that contained incorrect English which were given to pairs to revise.

Pairs in the first class (Class 1) self-transcribed their performances and followed essentially the same procedures in Lynch (2001). In Lesson 1, pairs performed their role-plays and then transcribed (verbatim) recordings of these performances. In Lesson 2, the pairs revised the verbatim transcripts for accuracy, and revised transcripts were collected by the teacher at the end of the lesson. In Lesson 3, the pair-revised transcripts students were returned, along with a teacher reformulated version of the same transcript. Pairs discussed differences between the two transcripts and with the teacher. Pairs then re-performed their role-plays at the end of the lesson. Three weeks later, pairs performed the same role-plays once again.

The Class 2 pairs followed the teacher-transcription procedure. In Lesson 1, pairs were recorded performing their role-plays, but not requested to transcribe these recordings. Rather, after the lesson, the teacher listened to the recordings himself and transcribed any portions of the recordings that contained incorrect English. In Lesson 2, each pair received a series of transcribed extracts and had to correct them, before checking revisions with the teacher. In Lesson 3, the pairs

reviewed the revisions, and then re-performed the role-plays. As with Class 1, pairs were performed the same role-plays once again three weeks later.

The study had two research aims. The first was to investigate whether the two TRT procedures were practical as routine classroom activities, and whether correcting learners transcripts or transcribing the recordings of role-plays imposed an excessive workload on the teacher. Lynch found that 90 minutes were sufficient for pairs to perform and then transcribe their role-plays, and the classrooms were large enough for pairs to listen to their own audiocassette players without being distracted by the others listening to theirs. The teacher preparation for either TRT procedure was reported as falling within the notional allowance at his institute, which was one hour's preparation time for a 90 minute lesson.

The second research aim was to “to see whether Class 1 learners’ active involvement in transcribing their performance would bring greater benefits than Class 2’s experience of reading transcribed extracts” (p. 316). To address this question, the role-plays performed three weeks after the TRT treatments were recorded and examined to see the extent to which the language items that had been revised on transcripts reappeared accurately. The pairs who had self-transcribed their performances achieved higher accuracy (64%) language items previously revised in repeated performance than the pairs who had not self-transcribed (47%). Lynch concluded that the difference was a result of differences in opportunities for noticing and language processing. His conclusion merits quotation at length.

The paired SI procedure in Class 1 created greater opportunities for both self- and peer-correction, as the partners were required to agree first on Transcript 1 [the verbatim transcript] and then on their corrections for Transcript 2. (p. 316)

This argument was plausible: the TRTs in Class 2 did not include a listen-and-transcribe stage, and therefore learners lacked the additional opportunities for noticing this stage could provide. Lynch, however, went on to write:

... the fact that the students worked on Transcripts 1 and 2 in pairs, not individually, meant they were encouraged to verbalize the process by which they were deciding how to improve their transcribed performance. This sort of interaction required them to engage in ‘language related episodes’—leading to the sort of co-constructed mental processing which can generate new L2 knowledge or consolidate existing partial L2 knowledge. *Although the data for my study did not include recordings of pairs’ discussion about form*, my earlier research with volunteers (Lynch 2001) did feature such interaction and found evidence of precisely this sort of co-constructed learning. (p. 317, emphasis added)

As the above shows, the 2007 study did not gather any data that could have supported Lynch’s conclusion and merely presumed that the Class 1 pairs had discussed language in the same way as the participants in his 2001 study. This presumption assumes that, when performing the same type of task, all learners will interact in similar manner, which—as discussed earlier in section 2.3—studies have repeatedly demonstrated not to be the case even when working on the exact same task. Furthermore, the pairs in Class 2 (teacher-transcription) did not work on transcripts “individually” during revision; rather, the procedures were “Each pair receives a series of transcribed extracts and has to correct them, before checking their changes with the teacher” (p. 312). However, the study did not (could not) say anything about the quality of the meta-talk in any pair- or pair-teacher discussions. To be able to conclude that student-transcription accounted for the differences in repeated role-play performance, the study would have to establish that, after the

pair-transcription stage, discussions in Class 1 during revision were comparable in quality to those in Class 2. Therefore, the only firm conclusion that can be drawn from the study is that, in Lynch's particular teaching situation, both transcript-revision-tasks were a practical means of focusing learner attention to form.

3.1.2 Mennim (2003, 2007, 2012)

Mennim's (2003) study was conducted in Japan with three university students with TOEFL scores of 500-550 taking a first year oral presentation course. In the course, the three students had worked as a group and had chosen a topic to research over a whole academic year (25 weekly classes of 90 minutes each). This group delivered an in-depth presentation of their research at the end of the course. The aim of the study was "to find out whether students could take advantage of a rehearsal of their final oral presentation in order to make improvements to their spoken output" (p.133).

Two weeks before their final presentations, the three students were audio-recorded delivering a 20 minute rehearsal of their presentation with the teacher as their only audience. As with their subsequent final presentations, students were only allowed use of small cue cards during rehearsal. After rehearsal, the group selected and typed a verbatim transcript of 5-minute segment that contained equal contribution from each group member. The transcript was printed and the group proceeded to make revisions in red pen. This revised transcript was taken away by the teacher who added corrections for errors they had missed. This teacher-revised transcript was returned to the students one week later (one week before the final presentations). The group's subsequent final presentation was audio recorded. The teacher transcribed the section of the final presentation that corresponded to the portion of the rehearsal the group had transcribed and revised. The noticing data collected were the group's repairs to the rehearsal transcript, the teacher's

corrections, and the excerpt from the final presentation. Mennim did not record the students' discussion as they corrected the transcript.

One finding was that students seemed to have used the transcription task for different ends. For example, one student appeared to have used the task much more than the others as an occasion to improve and expand on the content of her presentation (insofar as many of the revisions made to her portions of the transcript were comprised of insertions of additional clauses and other elaborations). In contrast, another student appeared to have focused exclusively on repairing language errors. Mennim speculated that the difference in proficiency between the two students (50 points in the TOEFL) might explain the difference.

As for group revision of errors, Mennim found that out of 122 errors he had identified in the rehearsal transcript, the students had identified 49 (40%), and that the majority were repaired accurately. The study then focused on three grammatical forms (articles, prepositions, and passive forms) which had received repeated teacher revision but which the group had infrequently noticed/repaired. Mennim reported that teacher feedback “seems to have been an important component of the task” (p. 137). For example, only two article errors had been identified by the group, whereas the teacher had identified 37 more occasions where articles were absent or used incorrectly. Of these 37 repairs to articles the teacher had suggested, 24 appeared in the final presentation of which 18 had been used accurately. Similar results were reported for errors in passive forms and prepositions. From these results, Mennim concluded the task had successfully focused student attention to form, and had led students to modify and improve their L2 output.

Mennim (2007) was conducted in the same university presentation course context as his earlier study. The study focused on two Japanese students and traced their acquisition of language form from the effects of noticing oral output. In this iteration of the course, TRTs were used as a regular classroom activity. While

pairs/small groups researched their topics over the course of the academic year, they delivered presentations in Months 3, 4, 7, and 8. After every presentation, each student transcribed a 5-minute portion of their group's presentation as individual homework. In the subsequent lesson, the groups would revise these transcripts. Both the students' presentations and interactions while discussing language problems were audio-recorded (p. 272).

The researcher examined the groups' noticing data (TRT discussions) and when an L2 form was noted or discussed, the researcher would trawl the presentation output data for subsequent re-emergences of that form. Only seven language forms fulfilled the researcher's criterion of appearing over 10 times in the presentation recordings were identified. The non-count noun 'garbage' was one of these. The study reported Mennim's tracking of the use of 'garbage' by two students, who researched and presented on the topic of garbage disposal.

The two students, Toru and Katsu, noticed the item three and five times over the year respectively while revising transcripts. In the first TRT, the pair was recorded questioning and deliberating whether 'garbage' could be pluralized. Although the pair appeared to lack the grammar terminology to describe this language point, they drew analogies to other nouns (such as 'information') and ultimately settled the issue by consulting their dictionary. In all subsequent TRTs, when erroneous use of 'garbages' was noticed, it was correctly revised without deliberation. Mennim found evidence that long-term gain in accuracy could be attributed to the repeated noticing of this form that occurred during the course. The word 'garbage' was produced by Toru 13 times over the three presentations and one rehearsal, while Katsu used it 25 times over the same period. The students went from an approximately 50% accurate use of 'garbage' in Months 3 and 4 to an accuracy rate of 100% four months later. The study, however, does not report

whether the other language forms that had also met the researcher's criterion for tracking-of-development had been similarly acquired.

Mennim (2012) returned to the data he collected in his 2007 study of students' discussions of language. In addition to Toru and Katsu's deliberation over the non-count noun 'garbage', the study presented an additional 7 episodes from the other six pairs/groups discussions to show how the students negotiated language form as they worked on their transcripts. The selected episodes illustrated the various strategies and resources students used to complete the tasks, which included their L1, dictionaries, metalinguistic knowledge, comparisons with existing forms they knew, formulating hypotheses about how L2 forms work, and subsequent re-application or refinement of these hypotheses (p. 55). Mennim noted that, "The length of some of these LREs indicates a sustained effort of cognitive processing such as hypothesis testing, generalizations, and use of metalanguage" (p. 60). He concluded by recommending student transcription/revision as "an effective way of generating discussion about language and encouraging learners to think about their own language use" (p. 61). The study did not, however, discuss how frequently the students had employed the strategies illustrated in the selected episodes.

3.1.3 Stillwell et al. (2010)

Stillwell *et al.* (2010) was conducted at a Japanese university with 20 freshmen students placed in the second highest English proficiency tier based on an in-house placement test. In the study, each student had made poster-presentations summarizing the main points of a challenging text on film genres. The students were divided into pairs and each student responded to questions about their poster posed by the other pair member for 3 minutes. The conversations, two for each pair, were recorded. After conversations, the pair transcribed both Q & A sessions (in class), and then worked together to produce corrected versions these transcripts.

Pair interactions during transcription and revision were *not* recorded. For homework, the students typed up the corrected transcript of the discussion their own poster only, and submitted it to the teacher for further correction. The teacher returned the transcripts the next week, and students then had three days to review the teacher feedback before repeating the entire task-cycle again with a new partner.

This was the only TRT study to investigate whether learners' transcriptions provided a faithful record of their actual performance. The researchers found students able to transcribe their work faithfully on the whole, but there were misrepresentations. Transcription errors included missing words, lexical chunks, and in some cases substantial portions of the recording missing from the transcripts. On average, pairs made 19 transcription errors for Presentation 1, but the numbers varied greatly among pairs with 38 being the most and 8 being the least. However, students overall appeared to become more adept at producing faithful transcripts. In Presentation 2, the average number of transcription errors decreased to 13, but again with great variation amongst pairs: a high of 39 for one pair and a low of 3 for another.

The researchers used Lynch's (2001) categories to catalogue the errors. On the whole, the class made a total of 301 corrections in the two presentations. Close to half (47.8%) of the pairs' corrections were for grammar, but only 1% of corrections were lexical. The remaining revisions were edits of dysfluencies (24.3%), reformulations (18.6%), and or 'mixed' (8.3%). Overall, pairs identified 52.35% of errors (301/575) produced in the two presentations. There was evidence that the students' ability to accurately identify and correct errors improved with practice, as the number of accurate corrections increased from 55 per cent in the first presentation to 60 per cent in the second, while the number of inaccurate

corrections dropped from 32 per cent to 20 per cent. However, the researchers noted pair performances varied widely.

The researchers also investigated whether, during repeat performance of the task, students had been “more likely to attempt reuse of self/peer-corrected forms or teacher corrections” (p. 447). As the quality of the Q & A poster sessions depended on the nature of the questions asked, Stillwell *et al.* reported “In many cases, the second presentation was significantly different to the first, and there was no attempt to reuse any of the corrections (teacher or student) the second time around” (p. 451). However, in cases where students “appeared to make conscious attempts to use the corrections”, the study found that 57 per cent of all the corrections remembered and used were teacher corrections. The researchers suggested that this higher uptake and reuse of teacher corrections was connected to students comments on the post-task questionnaires (see below) which indicated students possibly had more confidence in the native-speaker teachers’ ‘perfect’ English skills than their own ‘imperfect’ knowledge. However, this higher retention of teacher corrections is just as plausibly attributable to students having received this feedback just three days prior to repeated poster presentations, whereas student corrections had been made over a week before. The study further reported the accuracy of reuse was proportionally equal between student and teacher corrections (55% of each was used accurately), but the study also noted that some students reported that they had failed to review the corrections prior to the repeat performances (p. 448).

More positively, students appeared to have had a favourable impression of many aspects of the task. The post-task questionnaire asked students to rate the usefulness of each stage of the task cycle on a 5-point Likert scale (1 = not useful at all; 5 = very useful). Students valued most teacher correction (which 92% rated ‘very useful’) and self-correction (84% ‘very useful’), and also gave uniformly positive

ratings for transcription of one's own speech and repeating the speaking activity a second time. In contrast, the TRT stages involving the transcription and correction of a partner's speech were viewed the least favourably: rated 56% and 36% 'neutral' or 'not useful', respectively. The researchers again suggested that these latter trends may have been due to students' lack of confidence in their ability to revise each others' language errors correctly.

3.2 Discussion and summary of prior of research findings

There were a number of common elements to the studies reviewed above. All were designed with the classroom in mind so that the transcript revision tasks fit into existing university language course curricula. All learners were enrolled in EAP or first year courses with a focus on oral communication skills. Due to the pedagogic focus of these studies, they all describe their extant courses in some detail and how TRTs were implemented to supplement communicative tasks. Lynch (2007) addressed the amount of time required to include transcription within normal classroom activities, and the effect this had on teachers' work load. Both Lynch and Stillwell *et al.* (2010) reported that correction of transcripts was not onerous for teachers. All studies indicated that student transcription and revision could be completed in one to two, 90 minute lessons.

By making the students' L2 oral performance visible, TRTs appeared successful in promoting noticing. The studies reported learners on average able to identify 40%-60% of the errors they had produced; however, Lynch (2007) and Stillwell *et al.* (2010) found that learners failed to indentify the majority of their lexical errors. The studies also reported that the majority of errors learners identified were revised accurately, and Stillwell *et al.* found that students became more adept at revising error when TRTs were used repeatedly. Teacher corrective feedback (TCF) accounted for approximately 50% of revisions made to transcripts and TCF appeared to have been an important contributor to students' improvements in accuracy on repeated

task performance (Mennim, 2003, Stillwell 2010). Findings from Lynch (2007) and Mennim (2003, 2007) further indicated that gains in accuracy made by the students were durable. Taken together, the studies indicate the following: (i) TRTs are a practical means of having learners focus on form and error in their oral output; (ii) learners are able notice a considerable proportion of their errors and revise them correctly; and (iii) that learning benefits accrue from the transcript revision process.

While the studies were enlightening with regards to pedagogical practicalities and the revisions learners produced, they supplied remarkably little information as to how learners engaged with language during transcript revision. While Lynch (2001) reported that revisions were initiated almost equally between partners in 3 out of 4 pairs, no information was provided as to how the presence of a peer mediated the resolution of the language related episodes. Although Mennim (2012) provided examples of LREs which involved a range of deliberation strategies to solve problems, he did not indicate how frequently LRE resolution entailed such deliberation.

The complete absence of any investigation of the quality of attention learners paid to form left many previous TRT studies with little or no evidence to support their contentions. For example, Lynch (2007) concluded that having learners transcribe their errors prior to revision resulted in superior learning gains than when errors were transcribed by the teacher for learners to revise. His conclusion, however, was invalidated by the absence of any data which could demonstrate whether the students revising under the different conditions had discussed language to a comparable extent. Similarly, Mennim (2003) attributed differences in learners' uptake and retention of teacher feedback directed towards the same language forms to differences in the learners' L2 proficiency. However, there was no way to determine whether the extent of attention paid to the teacher feedback by the learners was comparable. Stillwell *et al.* (2010) contended that differences in the amount of 're-usage' of student-generated versus teacher-supplied revisions could be

attributed to learners valuing teacher-correction more, but this claim also was undermined by the absence of any examination of whether the attention paid to the two types of revisions was comparable. In sum, these studies seemed to assume that either all learners attended to language in a similar manner, or that any differences in quality of attention would not variably affect learning outcomes. Only one study (Mennim, 2007) traced the effect quality and extent of noticing had on learning outcomes, but was limited to discussion of just one pair of students, and one language feature.

A final limitation of the studies was that they largely failed to investigate students' perceptions of the TRTs as language learning opportunity. While the studies reported that student feedback was positive, the feedback elicited was focused exclusively on whether students found the task 'useful'. Stillwell *et al.* (2010) included results of a student survey conducted at the end of the study and reported that students had viewed "Correcting my partner's mistakes" as less useful than other aspects of the TRTs. While the researchers attributed this to students' lack of confidence in their ability to correct others' language errors, in the absence of student interviews, other explanations were possible. For example, perhaps learners simply meant that correcting the mistakes of others was not as directly beneficial for their own learning as correction of their own mistakes. In sum, the studies said little about what value learners saw in collaborating with others when revising, and what effect they felt the activity had on them.

3.2.1 Conclusion and expanded research questions

In conclusion, prior studies have examined TRTs predominantly in terms of what Breen (1987) would label their designer's 'workplan' or Coughlan and Duff (1994) as their 'behavioral blueprint' (i.e., the learning behaviours the teacher intends the task to elicit), but have provided little examination of how the learners actually performed and perceived the TRTs. While Lynch (2001, p. 131) noted that TRTs "offer a

productive route to noticing, in which learners are encouraged to externalize their thoughts about the formal correctness and semantic precision of their own output”, prior studies have given very limited examination to the actual use students made of these opportunities. Without a more thorough examination of learners’ engagement with language during transcript revision, the efficacy of TRTs as language learning and awareness raising activities cannot be adequately evaluated. To fill this research gap this thesis put forth the following expanded research questions.

1. What was the quality of learner EWL during transcript revision?
 - a) What were the foci, outcomes, and nature of the language related episodes (LREs) when pairs revised transcripts without teacher feedback?
 - b) What was the uptake outcomes and nature of LREs during transcript revision when pairs were provided with subsequent teacher corrective feedback?
2. Did differences in EWL during transcript revision affect retention of revisions?
3. How did learners perceive Transcript-Revision-Tasks as a language learning opportunity and experience?

4 CHAPTER 4: METHODS OF DATA COLLECTION

The next two chapters will describe the current study, where a combination action research and case study approach was chosen because the study examines phenomena—in this case ‘engagement with language’—within the real-life context in which it occurs (Duff, 2008). The current chapter will describe methods for data collection, and the next one will explain data analysis.

4.1 Action and case study research

Cohen and Manion (1994) state that *action research* is, “appropriate whenever specific knowledge is required for a specific problem in a specific situation; or when a new approach is to be grafted on an existing system” (p. 194). In this study, the new approach was the incorporation of Transcript Revision Tasks into a speech and debate course which previously had not provided opportunities for students to examine the linguistic accuracy of their oral performance. Action research involves “the systematic collection of data as planned interventions are enacted followed by analysis of what is revealed by the data and reflection on the implications of the findings for further action” (Burns, 2005, p. 59). In this study, the research purpose was to gauge how successful TRTs were at engaging my learners with language and how the tasks might be improved in that regard if re-employed in future iterations of the course.

Case study research is a widely-used research method in the context of second language learning, teaching, and use (Duff, 2008). Larsen-Freeman (1997) has noted that case study research has advantages for examining complex, non-linear systems like SLA, or in this study: engagement with language. Specifically, the characteristics of case study relevant for this study are boundedness and triangulation. Creswell explains the concept of boundedness by noting that a case study is an exploration of a “bounded system”; that is, a case or multiple cases (1998: 61). In this study, boundedness is relevant because I will describe peer editing

of transcripts in one L2 speaking skills classroom over a semester. In this study, each of the four pairs performing TRTs constituted an individual 'case' of learners engaging with language during peer editing (i.e., there were four cases in total). The concept of triangulation (Yin, 2003) refers to the researcher's examination of multiple sources of evidence to generate a more comprehensive perspective on the chosen cases. This study employed triangulation in both data collection (with multiple sources of data) and analysis (with multiple methods).

4.2 Research site and participants

The research was conducted at a Japanese national university which are considered to be, along with a few select private universities, the top-tier higher education institutions in Japan with concomitantly rigorous and competitive entrance requirements (Nagatomo, 2012). Students enrolled in the university's School of Language and Culture Studies are required to study two languages over their four-year programme: one language as their major field of study, the second as their minor. Therefore, students who enroll do so out of an interest in, and motivation to, study foreign languages. The majority of students who do not choose to major in English studies elect to minor in English.

4.2.1 The course (Interactive English: Speaking and listening)

The specific research site was a freshman English course taught during Japanese academic year 2014 (April 2014 to February 2015). The course, called *Interactive English: Oral production* (a pseudonym), is required for all 1st year students who minor in English. In the 1st year of the programme (but only in the 1st year) all students are streamed into one of four levels according to their scores on the *Test of English as a Foreign Language: Institutional Testing Program* (TOEFL ITP) which they are required to take before commencing the programme. The speaking and listening course I taught was intended for students with placement scores of 450-550.

The main goal of the course was for students to become more capable of making short presentations and holding in-class debate/discussions. As was stated previously in the introduction, the inclusion of Transcript Revision Tasks into the course was to address the shortcoming of previous iterations of the course where there had been almost no focus on the *correctness* of the language learners produced. The course under study was 30-weeks long with one 90 minute lesson per week, held over two, 15 week semesters. Piloting of certain aspects of procedures was conducted in Semester 1 (April-July 2014, inclusive) which primarily concerned providing learners with practice in self-transcribing and testing of equipment (described in Section 4.6.2). The Main Study was conducted in Semester 2 (October 2014 to February 2015, inclusive) and is described below.

4.2.2 *The participants*

There were eight students enrolled in the course, and all students agreed to and signed the consent form to participate in the study (see Appendix A). On the first day of the course, a student-profile questionnaire was given in class (see Appendix B). According to their profile questionnaires, students were all freshmen university students (18-20 years old). Four were males and four were female, and all were native speakers of Japanese. The average TOEFL IPT (*Institutional Testing Program*) score among the eight students was 479. As shown in Table 4.1 below, all students had received six years of compulsory English at junior high and high school, but two (Naoto and Chika) had additional English-learning years: in both cases while staying abroad in an English speaking environment. The names of all participants have been changed to protect privacy. Table 4.1 is organized by the pairs in which the students worked in the main study for the Transcript-Revision-Tasks. The *Interactive English: Oral production* class met once a week, but in addition these students were taking one other weekly required English course focusing on reading and writing skills (which I was not responsible for teaching). However, as their other English course

focused, in part, on the writing of argumentative essays, it was hoped students might find the two courses complimentary.

Table 4.1 Participants

Pair No.	Name	Age	Gender	TOEFL score	Years of English Study	Stay in English speaking environment
1	Naoto	21	M	520	9	3 years at English-medium international school while in China with family (ages 7-10)
	Ken	19	M	460	6	N/A
2	Chika	19	F	500	6	6 months study abroad at U.S. high school (aged 17)
	Momo	18	F	475	6	N/A
3	Naho	18	F	466	6	N/A
	Asami	19	F	458	6	N/A
4	Aki	19	M	480	6	N/A
	Yuta	18	M	473	6	N/A
<u>Key:</u> M=Male; F=Female; N/A= Not Applicable						

4.3 Data collection schedule and data sources

Under study was learner engagement with the Transcript Revision Tasks they performed in Semester 2 (Main study). Each TRT was situated within a larger 'Debate Cycle'. A single Debate-Cycle spanned three, 90 minute lessons (one 90 minute lesson each week of a given Cycle). A Debate-Cycle consisted of three inter-related classroom tasks. A brief description of this task sequence is outlined in

Figure 4.1 below (the stages of the Cycle are described more fully in subsequent sections).

Figure 4.1 Outline of lessons and tasks that comprised a 'Debate Cycle'

Lesson 1	
<u>Task 1: Initial delivery of position speech</u> • <i>Student A</i> (Agree position) delivers speech justifying his/her position. • <i>Student B</i> (Disagree position) delivers speech justifying his/her position.	
<u>HOMEWORK (Transcription)</u> Each student transcribes the recording of his/her delivery of position speech. Typed verbatim transcript is sent to teacher electronically (online).	
Lesson 2	
<u>Task 2: Transcript Revision</u> • Students taking the same debate position, collaborate in pairs to revise the language contained in the transcripts of speeches from LESSON 1. • Pairs compare their revisions to those made by the teacher.	
Lesson 3	
<u>Task 3: Redelivery (2nd delivery) of position speech + debate</u> • <i>Student A</i> (Agree position) redelivers position speech to a <i>new</i> Student B (i.e., <u>not</u> the same audience as LESSON 1.). • <i>Student B</i> (Disagree position) redelivers position speech. • Students proceed to debate by attempting to refute points from opponent's position speech and defend their own.	
<u>HOMEWORK (Transcription)</u> Each student transcribes the recording of his/her redelivery of position speech. Verbatim transcript of performance is sent to teacher electronically (online).	

The study was comprised three Debate Cycles, plus end-of-year teacher-student interviews held over the two weeks following the end of the final (3rd) Debate Cycle. The data collection schedule for the Main Study is shown in Figure 4.2.

Figure 4.2 Main study data collection schedule

Main Study (Semester 2)					
DEBATE CYCLE 1		DEBATE CYCLE 2		DEBATE CYCLE 3	STUDENT INTERVIEWS (pseudonyms)
Lesson 1 Speech 1a (27/10/2014)		Lesson 1 Speech 2a (01/12/2014)		Lesson 1 Speech 3a (19/01/2015)	Aki (05/02/2015) Ken (06/02/2015) Chika (09/02/2015)
Lesson 2 TRTs (04/11/2014)		Lesson 2 TRTs (08/12/2014)		Lesson 2 TRTs (26/01/2015)	Naho (10/02/2015) Naoto (12/02/2015) Asami (12/02/2015)
Lesson 3 Speech 1b (10/11/2014)		Lesson 3 Speech 2b (15/12/2014)		Lesson 3 Speech 3b (02/02/2015)	Yuta (13/02/2015) Momo (13/02/2015)

These three Debate-Cycles and the subsequent interviews generated the following data sources used to address the study's research questions: (1) audio-recordings and transcripts of students' initial delivery of speeches (24 transcripts), (2) audio-recordings of student-pairs' interaction during transcript revision (24 recordings), (3) transcripts revised by student-pairs (24 transcripts), (4) transcripts of students' redelivery of speeches (23 transcripts: one student was absent for one redelivery), and (5) recordings and transcripts of semi-structured student interviews (8 interviews).

4.4 Data collection and classroom procedures

The Debate Cycle procedures employed in the main study and the rationales for their design are presented below.

4.4.1 Debate Cycle, Lesson 1: Initial delivery of position speech

In the first lesson of a Debate Cycle, students were asked to prepare and deliver a speech explaining their reasons for supporting or opposing a debate resolution. These lessons were comprised of three stages, plus homework (self-transcription), described below.

Stage 1—Brainstorming ideas: From the start of the course, the four female students invariably chose to sit together in class, as did the four males. Throughout the main study, therefore, students were put into two brainstorming teams: Team 1 being all female and Team 2 being all male (see Table 4.2, below).

Table 4.2 Brainstorming teams for position speeches

Team 1 (female)		Team 2 (male)	
• Asami	• Naho	• Aki	• Yuta
• Chika	• Momo	• Ken	• Naoto

A team's goal was generate (brainstorm) ideas to support or oppose the Cycle's debate resolution. For the three Debate Cycles, the resolutions were:

1. "University students' summer vacation should be shorter"
2. "School uniforms should not be mandatory for junior-high school or high school students"
3. "All schools should be public: private schools should be abolished"

Once students were presented with a Cycle's debate resolution, one member from a brainstorming team played 'rock-paper-scissors' with a member from the other team; the team who 'won' could choose whether they would be arguing FOR or AGAINST the resolution. Both teams were then given 20 minutes to brainstorm together reasons to support of their position, and were asked to do so in *Japanese*.

Each team member had a brainstorming sheet on which to write—again in Japanese—the ideas generated by the group.

The rationale for having brainstorming in teams was that, in previous iterations of the course, I had found not all individual students able to generate at least three reasons for agreeing or disagreeing with a given resolution. The reasons for having teams brainstorm in their L1 (Japanese) was that previous experience lead me to suspect that brainstorming and writing notes in *English* could take longer than the 20 minutes allotted. My previous experience with Japanese students' 'brainstorming' was mirrored in Stone (2015, p. 13):

Students often carefully made these 'rough' notes with careful attention to the 'correctness' of their English meaning an activity that I would ideally have liked to take just a few minutes could quite often take twice as long. This hindered the purpose of the brainstorming, which was to quickly come up with a variety of ideas.

Stage 2—English *aide memoire* notes: Once the 20 minutes for brainstorming was up, team members were given an additional 20 minutes to sit and work individually to transform their Japanese brainstorming notes into brief English notes for use as an *aide memoire* during the student's delivery of their position speech. Dictionaries were allowed for this stage. These English notes were written on pre-formatted 'debate position houses' provided by the teacher designed to aid learners to organize and deliver speeches that followed the content-organization of position speeches suggested by their coursebook (Lubetsky, Lebeau, & Harrington, 1999). (See Appendix C for example of recommended content-organization of speeches). These 'debate houses' limited students to giving up to four reasons to support their position. Each reason itself could be further supported by up to three explanations/elaborations/or examples, and students were limited to writing up to eight words for any given support (see Figure 4.3, below).

The students were asked to make their English notes individually, rather than as part of team brainstorming, because in previous iterations of the course I had observed that when brainstorming, the specific English produced by the particular student who generated a given idea was almost always copied by all the other team members on their 'rough' notes. As this resulted in the English used on all students' debate houses being identical, and because it was desired in this study that the English used in a given student's position speech reflect that individual's choice of English to employ, the above procedures were adopted.

Stage 3—Delivery of speeches: Once students completed their English notes, each student selected a member from the opposing team as their partner. In these pairs, Student A ('agree') delivered his/her speech. Student B listened, took notes, and after the speech asked clarification questions to confirm understanding of opponent's speech if necessary. Student B ('disagree') then delivered his/her speech against the resolution. Both 'agree' and 'disagree' speeches were audio-recorded using MP3 digital recording devices.

Homework—Self-transcribing: The digital sound-files of the students' speeches were uploaded to the course *Moodle 2.6* website within three hours of performance. Across the three Debate Cycles, recordings of individual speeches were a median average of 7 minutes in length: the longest being just under nine minutes, the shortest just over 5 minutes (although this depended on fluency of delivery and whether student gave 3 versus 4 reasons to support position). The homework entailed each student download his/her sound-file, listen to their delivery of their own speech, and to transcribe what they heard verbatim.

Figure 4.3 Pre-formatted 'debate house' for position speech

RESOLUTION: _____ AGREE/DISAGREE			
<u>Reason 1</u>	<u>Reason 2</u>	<u>Reason 3</u>	<u>Reason 4</u>
<u>Supports</u> 1. (Eight words limit). 2. (Eight words limit). 3. (Eight words limit).	<u>Supports</u> 1. (Eight words limit). 2. (Eight words limit). 3. (Eight words limit).	<u>Supports</u> 1. (Eight words limit). 2. (Eight words limit). 3. (Eight words limit).	<u>Supports</u> 1. (Eight words limit). 2. (Eight words limit). 3. (Eight words limit).
NAME: _____, STUDENT ID _____, CLASS _____, DATE: _____			

Students had access to the freeware version of the transcription software *Express Scribe (version 5.13)*, and had been provided practice in transcription in Semester 1 (see, Section 4.6.2). No specific transcription conventions were prescribed other than (1) to indicate short pauses with two or three periods, and longer pauses with no more than ten; (2) to try to be consistent transcribing fillers such as 'ah', 'um' or 'mmm', including Japanese fillers such as 'eto' (i.e., 'um'); and (3) to type any other Japanese they spoke in Japanese script. Students sent their transcripts as 2.5 spaced word documents to the teacher within four days of end of lesson. Transcripts were sent using the course management system (CMS) software

Moodle (version 2.6). Transcripts were a median average of 403 words in length: the most being 629 words and the least 367 words (including fillers and dysfluencies such as false starts and repetitions). (An example of a student-transcribed speech can be found in Appendix D). A summary of procedures for Lesson 1 of a Debate Cycle is provided in Table 4.3, below.

Table 4.3 Lesson 1: Initial delivery of position speeches

Stage	Time	Activity
1. <i>Brainstorm</i>	20 min.	Student teams, speaking Japanese, brainstormed reasons in support of their debate position.
2. <i>Make speech notes</i>	20 min.	Students individually prepared English notes, in the form of ' <i>debate houses</i> ', for use during delivery of position speeches.
3. <i>Deliver speeches</i>	20 min.	Students formed pairs. Students delivered their position speeches. One student arguing in support of resolution; the other student arguing against.
(Homework)	Within 4 days	Students typed a verbatim transcript of their performance and sent to teacher.

4.4.2 Debate Cycle, Lesson 2: Transcript Revision

Lesson 2 of the Cycle was the transcription revision task (TRT) which had two stages: *initial pair-revision* followed by *final pair-revision*. Students paired with a revision-partner from their brainstorming team (i.e., who had also delivered a speech arguing the same side of the debate) they had self-selected in the first Debate Cycle. These pairings remained the same throughout the main study, and are shown in Table 4.4 below.

Table 4.4 Revision pairs

Team 1 (Female)		Team 2 (Male)	
<u>Pair</u> Chika & Momo	<u>Pair</u> Asami & Naho	<u>Pair</u> Ken & Naoto	<u>Pair</u> Aki & Yuta

Initial pair-revision of transcripts: Each pair sat in one of the four corners of the room, and students arranged their desks so that they were sitting face to face across from their revision-partner. The teacher gave each pair two hard copies of Student A's transcript, and two of Student B's. All transcripts had had line numbers inserted by the teacher the week previously. Pairs were also given MP3 digital recording devices and asked to press 'record' and leave recorders running throughout the remainder of the lesson.

One student then commenced to read his/her transcript aloud with the partner following along using their copy of the same transcript. Pairs were instructed to work together to make any revisions they felt could improve the formal accuracy or otherwise improve speeches. Students were allowed to consult dictionaries, but were not allowed to consult with the teacher. Any revisions the pair made to the transcript were made in pencil, and the revisions were to be written on both students' copies of the same transcript. Pairs were allowed 20 minutes to revise Student A's transcript; then an additional 20 minutes was given for the pair to revise Student B's transcript.

Final pair-revision of transcripts: After the 40 minutes for initial pair-revision, each pair was given two hard copies of each student's verbatim transcripts which had been revised by the teacher prior to the lesson. These teacher revisions had been typed in red font (teacher revisions are described in more detail next chapter). The teacher-revised transcripts had also had line numbers inserted which corresponded to those on the transcripts the pair had revised.

One student then commenced to reading aloud one of the *pair*-revised transcripts. The partner listened and followed along while looking at the corresponding *teacher*-revised transcript. When this student noticed a discrepancy in language between pair- and teacher-revised transcripts, he/she would interrupt the other student's reading and point out the discrepancy. Pairs had to decide what final revisions they wished to make in light of the teacher's revisions. Pairs could consult the teacher anytime during this stage. Any final revisions/changes students made were written in red pen. Pairs had 20 minutes to make the final revisions to transcripts of Student 'A's speech, and then 20 minutes for Student 'B's transcript. At the end of the lesson, all revised transcripts were collected by the teacher.

Students were asked, as far as possible, to perform transcript revision in English for two reasons. The first was that transcribing revision dialogue spoken entirely in Japanese was beyond my competence. The second was that, when applying for permission to conduct this study, my department head had felt that lessons spent discussing the English language in Japanese was not in-line with the institution's vision of what this English production course should be. A summary of the procedures for Lesson 2 of a Debate Cycle is provided in Table 4.5, below.

Table 4.5 Lesson 2: Transcript revision

Stage	Time	Activity
1. Initial revisions	20 min.	Student 'A' read his/her verbatim transcript aloud and worked with partner to revise language.
	20 min.	Student 'B' read his/her verbatim transcript aloud and worked with partner to revise language.
2. Final revisions	20 min.	Pair compared their revisions made to Student A's transcript to the teacher revisions made to same transcript to make final revisions.
	20 min.	Pair compared their revisions made to Student B's transcript to the teacher revisions made to same transcript.

4.4.3 Debate Cycle, Lesson 3: Redelivery of speeches and debate

This final lesson of a Debate Cycle lesson began with students sitting individually with only the following materials: (1) the 'debate house' *aide memoire* notes they had prepared two weeks prior, and (2) a blank, 'clean' copy of a debate house provided by the teacher. Students were given 15 minutes to make a revised debate house. That is, a debate house where the ideas to be presented were the same but where the English *aide memoire* notes may be reworded. Students did not have recourse to reading their revised transcripts and therefore had to rely on their learning and/or recall of their earlier transcript revisions.

Once updated debate houses were made, students selected a new partner of from the opposing side of the debate. That is, *not* the same person to whom they had delivered their initial speech performance to in Lesson 1. Students then sat in their pairs bringing with them only their new debate house and pencil. Pairs were given MP3 digital recording devices and asked to press 'record' and leave recorders running throughout the remainder of the lesson. Students then proceeded to deliver their speeches, with the listener taking notes on the back side of their debate house. Upon completion of redelivery of both the 'agree' and 'disagree' speeches, students then proceeded to debating (i.e., refuting and defending) the ideas the presented in the speeches.

The final homework of the Debate Cycle required students to download recordings of their redelivered speeches and debates, transcribe their own redelivered speech, and send to the teacher within 6 days. The pedagogic purpose here was to provide an opportunity for students to independently compare their two performances, and also provide them a record of their debate performance, to facilitate self-assessment, reflection, and informed goal setting. A summary of the procedures for Lesson 3 of a Debate Cycle is provided in Table 4.6, below.

Table 4.6 Lesson 3: Redelivery of speeches and debate

Stage	Time	Activity
1.	15 min.	Students worked individually to produce revised English aide memoire notes in the form of a fresh 'debate house'.
2.	60 min.	Students formed pairs with one student arguing in support of resolution and the other student arguing against (Note: <u>not</u> the same pairings as in Lesson 1). Students redelivered position speeches and proceeded to debate.
(homework)	Within 6 days	Students transcribed recordings of their redelivery of position speeches and sent to teacher.

The above Debate Cycle procedures generated data that was then analyzed (see Chapter 5) to examine the nature of pairs' engagement with language during TRTs, whether there were any differences in engagement between pairs, and whether these affected learning outcomes; namely, retention of revisions made.

4.5 Teacher-student interviews

After all three Debate Cycles were completed, I provided each participant with a one-on-one semi-structured post-task interview (see Appendix E for the interview guide). The aim of the interview was for the participants to reflect on their experiences of transcribing and peer-revising their position speeches. As research tools, however, interviews can have some shortcomings. Interviewees may not express what they really feel; perhaps to avoid embarrassment, or to provide answers which they think would please the interviewer (Wragg, 2002). Additionally, interviewees may also not reply in terms of their actual, but rather perceived, behaviour (Cohen, Manion, and Morrison, 2000).

To address the potential limitations above, an explanation was given prior to each interview and the student was told the purpose of the interview was to seek their assessment of self-transcribing and revising transcripts with a peer as language learning activities to help the teacher determine whether or not I would in the future employ TRTs in the debate course. This explanation of the purpose of the interview may have encouraged the students to speak truthfully since their responses would only affect my teaching practice but not affect their course grade nor directly require them to justify how they engaged with the tasks. Nonetheless, it was hoped that the individual student responses would assist in accounting for any differences in EWL between students and help shed light on factors which had facilitated or impeded engagement. While told that their participation in interviews was voluntary, all eight students agreed to be interviewed.

The interviews were conducted in English—with occasional code switching to Japanese—and lasted for 40 to 50 minutes each. Every effort was made to relax the interviewees. To reduce the ‘power-distance’ between teacher-interviewer and student, interviews were held on ‘neutral ground’, namely in the private small-group study rooms found in the university library. Before commencing, interviewees were offered bottled tea and snacks, and each was also asked whether they minded the interview being audio-recorded. There were no objections to being audio-recorded and so eight interviews were recorded for later analysis.

4.6 Transcription practice and piloting of procedures

Semester 1 of the course (the semester preceding the Main Study) provided students with two opportunities to practice transcription and one opportunity to perform a TRT.

4.6.1 *Transcription practice sessions*

The first transcription practice session was held in class in Week 5 of the semester. As every week, students had brought their laptop PCs to the lesson. At the start of

the lesson, students were given instructions about how to download Express Scribe Transcription software that allows users to slow down and loop sound files while transcribing, and convert or copy-and-paste their transcription into word documents. Once students had successfully downloaded and installed the freeware into their PCs (which took about 20 minutes), they were given handouts containing screenshots illustrating how to use the software and a teacher demonstration using a PC projector.

This lesson's pedagogic purpose was to be teaching students how to organize their ideas into a debate position speech (and into 'debate house' *aide memoire* notes). The coursebook provided a model position speech, and its recording (from the coursebook CD) was uploaded by the teacher on to the *Moodle* course management system prior to the lesson. The model speech was 222 words and 2 minutes in length. In the lesson, rather than play the coursebook recording to the whole class and/or have them read the example speech from coursebook, students sat side-by-side in pairs and were instructed to work together to (1) download the sound file, (2) listen to and transcribe the recording using the transcription software, and (3) convert or copy and paste transcript into a word document. Once completed, pairs could then check the fidelity of their transcription against the printed speech in their coursebook. With the teacher on hand to assist them with using the software, the entire practice session took approximately 50 minutes to complete. All pairs were able to produce near perfect transcripts of the recording with the exception of typographical (spelling/punctuation errors) differences. A similar pair-transcription practice session using the transcription software and recordings from the coursebook was held in Week 8 of Semester 1. Again, using the transcription software, pairs were able to produce perfect or near perfect verbatim transcripts of coursebook recordings.

4.6.2 *Piloting of Transcript Revision Task*

In Weeks 12-14 of Semester 1, students performed a Debate Cycle. This cycle differed from those of the Main Study in the following ways. At the end of the Week 12 lesson, student teams had brainstormed (in Japanese) ideas for or against the resolution “Arranged marriages are better than love-based marriages” to be debated the following week. As homework, students prepared their individual ‘debate house’ *aide memoire* notes.

In Week 13, student pairs delivered their position speeches and debated. These speeches and debates were recorded using MP3 digital recorders. For homework, students were instructed to individually transcribe verbatim their own speech delivery and send transcript to the teacher. These speeches/transcripts were of comparable length to those produced in the Main Study.

In Week 14, student pairs were given two hardcopies each of the transcripts to revise. As in the Main Study, one student would read his/her transcript aloud while the partner would follow along while reading their copy of the transcript. However, unlike in the Main Study, pairs were not provided with teacher-revised transcripts. Instead, the teacher attempted to provide corrective feedback (CF) directed towards the revisions student-pairs had made by collecting and revising each transcript as it was finished being revised by a given student pair. Two immediate findings from this piloted TRT were:

- Pairs took approximately 15-25 minutes to finish revising a transcript (which, in part, depended on its length).
- It was logistically impossible for the teacher in class to read and provide teacher CF to all of the pair-revised transcripts.

Therefore, in the Main Study the teacher made revisions to students’ transcripts prior to TRTs and provided these to the pairs after they had finished revising

transcripts. As piloting indicated pairs would take approximately 20 minutes to revise a transcript, this was the time limit allotted in the study.

During the summer break between semesters, when I began listening to recordings of pair interaction and reading their revisions in an attempt to match them (and begin to develop the study's methods of data analysis), the following problem was found. During the pilot TRT, pairs had been joined their desks together and sat side-by-side. It was realized that when sitting this way, students would often lean over and physically point at their partner's copy of a transcript to draw partner's attention to a linguistic feature they felt in need of revision. This produced recorded exchanges such as the example below. (Note: ALL CAPS indicate students reading transcripts verbatim; 'apostrophes' indicate the revision students are suggesting—see Appendix F for transcription conventions of interaction data).

Momo: AND IF SUCH GOOD PERSON CHOOSES ME, MY PARENTS FEEL GOOD.

Chika: We need 'a' here. And we missed back here. This needs '-ed' I think, because you say about something you doing before.

While it could be assumed from the pair-revised transcript that Chika was recommending the revision *such a good person*, it was not possible to tell where 'back here' in the transcript Chika was recommending adding an '-ed' to (presumably) render a verb into the past tense. When listening to recordings, in many cases it was impossible to match the revision a pair had made to the pair's language related episode. Therefore, to collect student-interaction data more amenable to analysis, in the main study:

- To reduce incidences of students physically pointing to language features on the transcript lying in front of their partner, student pairs were arranged so that they were sitting face to face across their desks.

- To make it easier for students to verbally refer to specific language features contained in transcripts, both the unrevised and teacher-revised transcripts given to student pairs had each line of the transcripts numbered, with corresponding line numbers on both transcripts.

5 CHAPTER 5: DATA ANALYSIS METHODS

5.1 Fidelity of student self-transcription

Students had been instructed to transcribe their initial delivery of speeches and subsequent re-delivery of the speeches word for word. Practice in using transcription software had been provided in the previous semester. Both transcripts were, of course, meant to include their errors. In the case of the initial speeches, transcribed errors were meant to be discussed and corrected in pairs. The purpose of having students transcribe verbatim the re-delivery of their speeches was so that a student could track their own language development. At issue here was whether students had provided faithful records of their actual speech performance, or whether they appeared to be ‘cleaning up’ their transcripts by self-revising errors, perhaps to avoid embarrassment, prior to submitting their transcription to the teacher or showing it to a peer during transcript revision tasks.

To confirm the fidelity of student-transcription, I listened repeatedly to the recordings of the speech deliveries while simultaneously reading the student-produced transcripts and made note of any discrepancies. I found that students had included almost all the errors that appeared in the actual speeches, and showed a clear attempt to record their output faithfully. This was evident from the many dysfluencies (hesitations, pauses, repetitions, false starts, and outright ‘breakdowns’) they included in their transcripts. Unlike Stillwell *et al.* (2010), when comparing audio recordings to the student-made transcriptions, no substantial portions (‘chunks’) of oral performances missing from transcripts, which I ascribe to students having been previously provided with transcription freeware and practice. Also, no cases were found of students adding new sentences or ideas into the transcripts.

However, a small number of minor differences were found between what I heard and what students had transcribed. The first type of discrepancy appeared to

be deletions of minor dysfluencies or self-corrections that had been produced during speech delivery, as in the examples below:

Actual utterance: The number of absent students **have has** become more.

Student transcription: The number of absent students **has** become more.

Actual utterance: I **have been... went** to a study abroad this summer.

Student transcription: I **went** to a study abroad this summer.

These examples may be evidence of students deliberately ‘tidying’ speeches during self-transcription. However, in this study, even when such self-repairs were transcribed and the initial inaccuracy subsequently excised by pairs from transcripts, such revisions were never counted as ‘corrective’ (see *section 5.2.1* below) as they represented the mere deletion an already self-repaired form. Therefore, these types of discrepancies had no impact on the study.

After cases of failures to transcribe dysfluencies were discounted, across 47 transcripts (of both initial and redelivered speeches) of on average just under 400 words (including transcribed dysfluencies), students made a median average of only 12 ‘misrepresentations’ on transcripts, with 18 being the most and 6 being the least. This was far less than Stillwell *et al.*’s (2010) learners, which again I attribute to this study’s provision of transcription software and practice to students.

These misrepresentations on transcripts, to my ear, either (a) rendered inaccurate utterances accurate or (b) made accurate utterances appear inaccurate on the transcript. Errors produced orally which appeared to have been ‘corrected’ during student transcription, were all minor improvements such as insertion of function words (e.g., articles, auxiliary verbs, conjunctions, or prepositions) or lexemes, for example: adding ‘s’ so that noun and verbs agreed. I found no examples major improvements being made to inaccurate utterances, such as the insertion of missing nouns or verbs, changing of word class, replacing one word with another, or reformulation of clauses. When self-transcribing their *initial* delivery of speeches,

any corrections students made to oral errors were, obviously, never revised by pairs during TRTs, and therefore had no bearing on this study's analysis of engagement and retention of revisions.

However, depending on the student, 40-60% of misrepresentations found had, to my ear, made the student's oral speech delivery appear less accurate on transcripts. These included cases such as:

- Single words said but simply not transcribed, these were usually articles and prepositions, but also included nouns and verbs.
- Obvious typing/spelling mistakes, for example: transcribing, "All parents expect their children..." as "All parents except their children..." or "Although" as "also".
- Not transcribing the 's' at the end of a word when the sound elided with the 's' at the beginning of the next word, for example, transcribing what I heard as "All teachers should" as "All teacher should".

Cases of transcribed 'errors' contained in transcripts that I could establish had never been made or I was unsure had actually been uttered were 'flagged'. Pair-revisions of such misrepresentations were discounted when re-deliveries of speeches were analysed for signs of retention of revisions (Section 5.5, below). Additionally, before starting the analysis of transcripts for language development, I had to ensure that transcripts of students' *redelivery* of speeches were a 100% accurate reflection of actual performance. I amended the student-made transcripts whenever I found discrepancies between transcripts and audio-recordings of performance.

5.2 Identifying LREs in pair interaction

As requested, pairs mainly used their L2 (English) during transcript revision. Over the course of the study, each of the four pairs made revisions to six transcripts of student speeches (3 speeches produced by Student 'A' and 3 by Student 'B'). Each transcript received two rounds of revision: in Round 1, pairs made initial revisions

without recourse to teacher feedback; in Round 2, pairs could compare their revisions to those made by the teacher to make final revisions.

I transcribed verbatim the 24 protocols of 20-minute audio-recordings of pair interaction during initial revision (Round 1), and the 24 protocols of 20-minute audio-recordings of pair interaction during final revision (Round 2) (48 protocols in total). After transcribing, to investigate the nature of pair interaction and how languaging mediated learning opportunities during transcript revision, I used language related episodes (LRE) as a unit of analysis. An LRE was defined in this study as any portion of pair-talk in which students discussed making a modification to an original transcript of a delivery of a debate position speech.

To identify the LREs contained in the pair-talk, firstly sentences from the original (unrevised) transcripts were matched with the same sentences as modified by the pair, and the matching sentences were placed into a table. Next, pair-talk where students had discussed the changes made to the sentences was copied from the transcripts of student interaction and pasted into the table. LREs were then identified and matched to revisions. Following this study's transcription conventions (see Appendix F); language taken from speech-delivery transcripts was shown in ALL CAPS. As pairs' initial revisions had been made in pencil, these were replicated in *handwriting* font, underlined, and numbered. An example of an LRE matched to a revision is shown in Figure 5.1, below:

Figure 5.1 Example of LRE matched to revision

Original [Line 3]	WHEN WE HAVE SCHOOL UNIFORMS WE DON'T TO BUY MANY CLOTHES TO WEAR ON WEEKDAYS.	
Initial Revisions	WHEN WE HAVE SCHOOL UNIFORMS WE DON'T (1) <u>NEED</u> TO BUY MANY CLOTHES TO WEAR ON WEEKDAYS.	
Pair-talk: Initial revision (A = Asami [author]; N = Naho [revision partner])		
12	A: WHEN WE HAVE SCHOOL UNIFORMS WE DON'T... TO=	LRE 1
13	N: We don't to?	
14	A: We don't ' <u>need</u> ' to,	
15	N: 'need'	
16	A: WE DON'T NEED TO BUY MANY CLOTHES.	

In Figure 5.1 above, turns 12-16 were counted as a single LRE. Identifying this LRE was straightforward as it was an example of what Fortune and Thorp (2001) label a *continuous* LRE; that is, learners discussed a single language point and concluded the discussion without returning to the form later.

In contrast, Figure 5.2 below provides an example of where it was more challenging to isolate the four LREs contained in the pair-talk. Below, the pair is revising Yuta's (the 'author') speech transcript. The excerpt illustrates two examples of *discontinuous* LREs (Fortune and Thorp, 2001) where the speakers leave a point and return to it later. For example, in turn 9, Aki first proposes inserting 'will' between the words 'it' and 'look', a suggestion Yuta hesitates to accept. Aki then leaves the point and moves on. But later, Aki reiterates this suggested revision in turn 17 and the potential revision is discussed until turn 22 when the revision is rejected by Yuta. Therefore, turn 9 was identified as the first portion of LRE 2 (labelled LRE 2a) and turns 17-22 as LRE 2b (the continuation of the LRE). Because the pair-talk contained discontinuous LREs, it also contained examples of *embedded* LREs which were, 'necessarily preceded and followed by a discontinuous one' (ibid,

p. 156). The examples below serve to illustrate the care that was required when identifying and segmenting individual LREs.

Figure 5.2 Examples of discontinuous and embedded LREs

Original [Line 10-12]	SO IF THOSE PEOPLE WEAR SCHOOL UNIFORMS IT LOOKS VERY BAD OR VERY TERRIBLE, SO SCHOOL UNIFORMS SHOULD NOT BE MANDATORY.	
Initial Revisions	SO IF THOSE (1) <u>STUDENTS</u> WEAR SCHOOL UNIFORMS, (2) <u>THEY LOOK</u> VERY BAD OR VERY TERRIBLE. (3) <u>IN SUCH WAY</u> SCHOOL UNIFORMS SHOULD NOT BE MANDATORY.	
Pair-talk: Initial revision (TRT 2) (Y = Yuta [author]; A = Aki [revision partner])		
7	A: IF THOSE PEOPLE...those 'students' kana? (<i>isn't it?</i>)	LRE 1
8	Y: Ok. (writes down and reads suggested revision) IF THOSE STUDENTS WEAR SCHOOL UNIFORMS IT LOOKS VERY BAD.	
9	A: 'It will' kana (<i>shouldn't it be?</i>)? It ' <u>will</u> ' look bad.	LRE 2a
10	Y: Hmm (unsure) <6>	
11	A: Umm...and just this part...is difficult to understand 'so'. I think could be better change 'so', but how to change...I don't know...how to change?	LRE 3a
12	Y: Hmm (unsure) <20>	
13	A: For example, this 'so' and this 'so'...	
14	Y: Double 'so' is strange? Ok, change second 'so' to 'in such way'.	
15	A: Hmm (unsure) <6> but here...It is...IF THOSE STUDENTS WEAR SCHOOL UNIFORMS IT...if those students wear school uniforms 'they' looks very bad.	LRE 4
16	Y: 'They.' Ok. 'They <u>look</u> '...no 's'.	
17	A: 'They will look'.	LRE 2b
18	Y: ' <u>Will</u> '?	
19	A: 'If' <i>desu kara</i> (<i>because of 'If'</i>)	
20	Y: Where?	
21	A: Starts 'If' so need 'will'.	
22	Y: I don't know...'they look very bad' I think. It is correct. I stay 'they look very bad'. <6> Is this difficult to understand this, uh, content?	
23	A: Contents is not difficult but the...uh...how to go to...to connection the next sentence...uh there is some problem...this part is not natural to go to the next sentence...'so' 'so' 'so' always 'so'.	LRE 3b
24	Y: mm (unsure).	
25	A: So I want to change the sentence, the sentence be more natural...more sophisticated.	
26	Y: But 'so' 'so' is wrong?	
27	A: Um...not wrong <7>but not natural.	
28	Y: Ok, this 'so' be changed 'in such way'.	

29	A: I'm not <10> sounds something wrong.	
29	Y: No, I think no. I will change 'in such way'.	
30	A: I don't think..., uh but ok, change.	

5.2.1 Matching LREs to revisions

All of the above LRE examples were concerned with making changes or *language amendments* intended to correct or stylistically improve upon the language already contained in verbatim transcripts. In this study, a language amendment and a 'revision' are used interchangeably. However, the students also made another type of modification to transcripts I did not consider to be revisions. These modifications were *deletions of dysfluencies*, where pairs would strike-through (i.e., delete) dysfluencies such as false starts, repetitions, fillers, and hesitations. These were naturally occurring features of speech (although students were often more dysfluent than a competent L2 speaker), but not language features learners were trying to improve or repair. While language revisions could always be matched to an LRE, frequently this was not the case for deletions of dysfluencies. The example in Figure 5.3 illustrates the problem.

Figure 5.3 Example deletion of dysfluencies

Original [Line 24]	SO I THINK IT IS IMPORTANT....I THINK THEIR FEELINGS ARE MOST IMPORTANT.
Revision:	SO I THINK IT IS IMPORTANT.... I THINK THEIR FEELINGS ARE MOST IMPORTANT.
Pair-talk: Initial revision (K = Ken [author]; N = Naoto [revision partner])	
63	K: SO I THINK IT IS IMPORTANT I THINK THEIR FEELINGS eh? (laughs) <8> Ok?
64	N: (grunts)
65	K: SO I THINK THEIR FEELINGS ARE MOST IMPORTANT.

It was inferable from the above that Ken had stopped reading aloud to delete his false start, and had showed his deletion to his partner (“Ok?”) who had made the same deletion on his copy of the same transcript. However, as there was no explicit verbal reference to what had been deleted this was not considered an LRE. Students would also commonly just skip over dysfluencies such as fillers and hesitations when deleting them and not read them aloud. Sometimes students would explicitly refer to a deletion (e.g., “SO, UH BUT I THINK, ah, cut ‘so’ ‘uh’”) or make meta-comments about dysfluencies (e.g., “I say ‘um’ so much”); while such LREs were counted, they were not analysed further nor included in this study’s analysis of LREs.

5.3 Analysis and coding of LREs from initial revision pair-talk

The LREs from recordings of student interaction during initial revision without teacher feedback (Round 1) and the LREs produced during final revision with teacher feedback (Round 2) had somewhat different coding procedures. This section explains the coding procedures for LREs produced during initial pair-revision, and Section 5.4 explains the coding procedures for LREs produced during final pair-revision of transcripts. LREs produced during initial pair-revision were coded for (1) focus, (2) outcome, (3) quality of metatalk, (4) source of initiation, and (5) source of resolution.

5.3.1 LRE foci

To investigate the aspects of language learners focused on revising, LREs were first analysed for their foci. The LREs in this study fell into three broad types. The first was a *form focused LRE* (F-LRE) in which a pair discussed revising an discrete syntactic or morphological feature, such as: verb morphology, auxiliary verbs, noun plurals, pronouns, the genitive ‘s’, word class (e.g., equal vs. equality), word order (e.g., ‘get money enough’ vs. ‘get enough money’) and insertions or changes to determiners and conjunctions. Excerpt 1 below provides an example of an LRE focused on verb morphology (subject-verb agreement).

Excerpt 1—Form focused LRE (F-LRE)

- 6 Ken AND SCHOOL UNIFORMS DOESN'T BECOME OR SUIT
EVERYBODY.
7 Naoto This is plural, so 'don't'.
8 Ken 'Don't'

Revision: AND SCHOOL UNIFORMS DON'T BECOME OR SUIT EVERYBODY.

The second type of LRE was a *lexis focused LRE* (coded L-LRE) which dealt with word meaning, word choice and/or phrase choice (e.g., 'On the other hand' vs. 'In the opposite way') and choice of prepositions. Excerpt 2 provides an example.

Excerpt 2—Lexis focused LRE (L-LRE)

- 79 Chika SCHOOL UNIFORMS CAN BE WORN FOR THE MARRIAGE
CEREMONIES AND DEAD CEREMONIES...Something wrong
I think.
80 Momo Students can wear to these, these ceremonies.
81 Chika I understand what you want to say, but...umm...ceremony
is not 'dead'.
82 Momo Person ... people dead.
83 Chika (laughs) I know!
84 Momo (laughs) I want to say 'Sōgi' (*funeral*).
85 Chika I know but <25> (checking dictionary)
86 Momo You find it?
87 Chika Minute <10> 'Sōgi' 'funeral'.
87 Momo Fu?
88 Chika 'Funeral' f-u-n-e-r-a-l.
89 Momo *Kashitte (lend me/show me [the dictionary]).*

Revision: [...] MARRIAGE CEREMONIES AND FUNERAL CEREMONIES.

The third type of LRE focused on making a *reformulation* (REFORM) where students deleted a sentence or clause entirely and rewrote it from scratch. This type of LRE did not focus on a single clearly identifiable language feature. That is, pairs sometimes appeared to have felt discrete revisions insufficient for repairing the language of the original speech, and that in order to convey the intended meaning clearly, the original wording had to be largely scrapped. Reformulations were

counted by the clause, with one clause representing one reformulation. An example of a reformulation focused LRE is shown in Excerpt 3.

Excerpt 3—Reformulation focused LRE (REFORM)

- 38 Chika SO IT'S VERY NECESSARY FOR THE UNIVERSITY STUDENTS
MAKING MONEY DURING THE SUMMER VACATION... Ah, this is
not good, I want to say, you know, for students=
39 Momo For university students.
Chika Eh? ... yeah, for university students. But I wanted to say here for
long vacation to do things like I say here to have a long vacation
making money...oh, change this. I mean *Gakusei wa okane o*
kasegu tame ni natsuyasumi ga hitsuyōdesu (*Students need*
summer vacation to earn income) I wanted to say it.
40 Momo Oh.
41 Chika But during my speech I couldn't say it.
42 Momo You tried to say.
43 Chika Yeah. So <9> IT IS NECESSARY FOR THE UNIVERSITY STUDENTS
<23> (sounds of writing) 'to have long summer vacation'=
44 Momo *Matte (wait)* (sounds of writing) <16> ok, ok.
45 Chika 'to get the' no, uh <8> 'to make money'. So let's change this like
this <20> (sounds of writing)

Revision: SO IT'S VERY NECESSARY FOR THE UNIVERSITY STUDENTS TO
HAVE THE LONG SUMMER VACATION TO MAKE MONEY.

5.3.2 LRE outcomes

To discover how each pair resolved their discussions of language, LREs were coded for their linguistic outcomes. Outcomes were coded as (a) Positive, (b) Negative, or (c) Unresolved.

(a) Positive outcomes (-/+) (+/+)

There were two types of positive LRE outcome. The first was when a LRE resolution rendered inaccurate language into something linguistically accurate. Such outcomes were coded (-/+). The second type of positive outcome was where something already accurate was replaced by an equally acceptable alternative.

These were coded as (+/+). The following excerpts provide examples of both types of positive outcome.

Excerpt 4—Positive outcome: (-/+)

- 23 Naho WE CAN HAVE A LOT OF EXPERIENCE IN ONLY ONE
VACATION ...'experiences'
24 Asami Mm (yes) 'experiences'.

Revision: WE CAN HAVE A LOT OF EXPERIENCES IN ONLY ONE VACATION.

Excerpt 5—Positive outcome: (+/+)

- 28 Naho AND WHEN WE START WORK AFTER GRADUATE
UNIVERSITY...Ah, this sentence need 'start working'.
29 Asami Ah.

Revision: AND WHEN WE START WORKING AFTER GRADUATE UNIVERSITY

(b) Negative outcomes (-/-) (+/-)

There were two types of negative LRE outcome. The first rendered something originally inaccurate into something also inaccurate (-/-). The second replaced something already accurate with something inaccurate (+/-). The excerpts below show each type of negative outcome.

Excerpt 6—Negative outcome: (-/-)

- 14 Ken I see... BECAUSE STUDENTS WANT TO GO TRIP. Students
want to 'go to trip?'
15 Naoto 'Go trip', go to trip, go trip, go to trip (sounding
alternatives aloud) which is better?
16 Ken (Clears throat) <7>Hmm 'go to trip'
17 Naoto 'go to trip?'
18 Ken I think so, trip you go to Disney, go to Hawaii, like this.

Revision: [...] BECAUSE STUDENTS WANT TO GO TO TRIP.

Excerpt 7—Negative outcome: (+/-)

- 26 Yuta But 'so' 'so' is wrong?

- 27 Aki Um...not wrong <7>but not natural.
 28 Yuta Hm...ok, this 'so' be changed 'in such way'.
 29 Aki I'm not <10> sounds something wrong.
 29 Yuta No, I think no. I will change 'in such way'.
 30 Aki I don't think..., uh but ok, change.

Revision: SO IF [...]. IN SUCH WAY SCHOOL UNIFORMS SHOULD NOT BE MANDATORY.

Excerpt 7 illustrates that, in this study, mutual agreement over a revision was *not* necessary for an LRE to be considered resolved. In Excerpt 7, Yuta—who had delivered the speech—exercised what Villamil and De Guerrero (1996) would call his ‘authorial control’ and insisted upon making the revision to his transcript despite the reservations expressed by his partner. *Reformulations* were coded being a positive outcome only if they were error free. As far as possible, I restricted myself to judging LREs as having ‘negative’ outcomes when they lead to production of an *absolute error* (linguistic forms that violate the TL ‘code’ and are not possible under TL rules) rather than merely a *dispreferred form* (forms that are possible but arguably less natural) as the latter involve more subjective judgments of acceptability (Ellis and Barkhuizen, 2005, p. 59). However, I acknowledge that this distinction is continuous rather than absolute and at least partly subjective.

(c) *Unresolved LREs (?)*

An LRE coded as unresolved (?) was one where a pair could not solve a language problem, abandoned resolving the language point, and made no revision; or a pair could not choose between two proposed revisions and both were written on the transcript. In cases where discussion over revision of a language feature resulted in an unresolved LRE, the researcher underlined and annotated the feature with a question mark [?]. The excerpt below illustrates an unresolved (abandoned) LRE.

Excerpt 8—Unresolved LRE: (?)

- 44 Yuta IN LONG VACATION PEOPLE WILL TRY TO DO DANGEROUS THINGS. FOR EXAMPLE TRY TO DO DRUGS AND SOME WILL SOME CRIMES SO IF SUMMER VACATION IS TOO LONG PEOPLE BECOME VERY LAZY.
- 45 Aki Um, problem <9> Drugs and some crimes is not...does not come from lazy.
- 46 Yuta 'Lazy' is wrong?
- 47 Aki Match, doesn't match, uh <12> what you said example is not 'lazy' people. 'Lazy' is wrong.
- 48 Yuta Change only one word is difficult. Hmm <45> you know?
- 49 Aki To change? No, maybe no.
- 50 Yuta *Tsugi (next/move on).*

(Non) Revision: [...] PEOPLE BECOME VERY LAZY [?].

Here it needs mention that not every LRE that was resolved lead to the writing of a revision. In a few cases, there were LREs in which a pair had discussed *potentially* making a revision, but ultimately decided that the language feature was not in need of revision. Such LREs were counted and coded. The distinction between an *unresolved* LRE and a *resolution not to revise* was felt to be important as they represented very different outcomes: a correct decision not to revise was coded as having a positive outcome (+/+), and an incorrect decision coded as a negative outcome (-/-).

5.3.3 Quality of metatalk

To begin investigating students' engagement with language during revision, LREs were coded for quality of metatalk. Quality of metatalk referred to the extent to which a pair discussed a language point, and provided a window into the degree to which learners were cognitively engaged with the revision tasks. Based on the distinction drawn in Storch (2008) and Storch and Wigglesworth (2010), an LRE was coded as either showing *limited* (L) or *extended* (E) metatalk.

Limited (L) metatalk referred to LREs where one student (a) suggested a revision without explanation/justification and (b) the suggestion was accepted by the partner without further discussion. Excerpt 9 illustrates limited (L) metatalk.

Excerpt 9—Limited (L) metatalk

- 51 Momo TWO MONTH IS ENOUGH TIME I THINK=
52 Chika Two 'months'.
53 Momo Ah! yes, yes <4>

Revision: TWO (1)MONTHS IS ENOUGH TIME I THINK [...]

Extended (E) metatalk referred to pair discussion where either (a) a suggested revision was questioned or an explanation was provided or sought; and/or (b) alternative revisions were suggested. Excerpt 10 illustrates extended metatalk.

Excerpt 10—Extended (E) metatalk

- 53 Momo <4> TO TRAVEL ABROAD, RETURN OUR FAMILY...return
'to'?
54 Chika Need 'to' but I think this part is weird, how about 'go back
to' our family or ...
55 Momo Not correct?
56 Chika 'Visit!', 'visit', yes. You mean visiting family.
57 Momo But...it has same ... is same as my idea 'return'. Different?

58 Chika Uh, because you will just visit and come back again, right?
Come back to Tokyo again=
59 Momo Yes.
60 Chika So I think it is, just like, short visit in vacation with your
family. 'Visit' I think is better. You won't return and like uh
be living with them.
61 Momo Hmm <20>
62 Chika Yes? Change it?
63 Momo Ok. 'Visit' <11> (writes 'visit') with not 'to'.

Revision: [...] TO TRAVEL ABROAD, VISIT OUR FAMILY [...]

5.3.4 Coding for LRE initiation and resolution

To investigate the degree to which a member of a pair participated in the revision of transcripts, LREs were coded for their *source of initiation* and *source of resolution*. Following Lynch (2001), 'initiation' referred to which member of the pair first noticed and explicitly mentioned the language feature perceived to be in need of revision. LREs started by the student whose speech transcript was being corrected (i.e., started by the speech's 'author') were coded as having been *self-initiated* (SI). Discussions of revisions started by the revision partner were coded as having been *other-initiated* (OI).

'Resolution' referred to which member of the pair supplied the revision. LREs were coded as having been (i) *self-resolved* (SR) by the speech author; (ii) *other-resolved* (OR) by the revision partner; or (iii) *co-resolved*.

Excerpt 11, below contains three example LREs to illustrate self-/other- LRE initiation and resolution. In LRE 1 (turn 88), Ken—the speech 'author'—questioned his use of 'protest', and Naoto—the revision partner—supplied the revision 'protested' (turn 89). Although Naoto provided the resolution in a questioning tone, Ken's contribution to the LRE resolution was merely to affirm his acceptance of Naoto's proposed resolution. Therefore, LRE 1 was coded as being self-initiated by Ken and other-resolved by Naoto (SI/OR). In LRE 2 (turn 90), Ken noticed that the verb 'was' needed insertion in front of 'so cold', and did so: LRE 2 was coded self-initiated and self-resolved (SI/SR). In LRE 3 (turn 95), Naoto noticed and corrected Ken's use of 'get off the sweater', and this LRE was coded as other-initiated/other-resolved (OI/OR).

Excerpt 11—Examples of self- and other- LRE initiation and resolution

- | | | |
|----|-------|--|
| 88 | Ken | I PROTEST HIM BECAUSE IT SO COLD. 'Protest'? |
| 89 | Naoto | 'Protest', 'protest'. What is the past of 'protest'...uh, 'protested'? |
| 90 | Ken | Hm (affirms). I PROTESTED BECAUSE IT 'was' SO COLD. |

- 91 Naoto Wait, (writing revision) 'protested' and 'it was' SO COLD.
 92 Ken BUT I WAS FORCED TO GET OFF THE SWEATER.
 93 Naoto The sweater?
 94 Ken Yes, it was not uniform.
 95 Naoto 'Get off' *chigau* (wrong) 'take off' I think. Let me <50>
 (checking dictionary) yeah, needs 'take off'.

Revision: I PROTESTED HIM BECAUSE IT WAS SO COLD BUT I WAS FORCED
 TO TAKE OFF THE SWEATER.

Co-resolved (CoR) LREs were those in which both pair members contributed to arriving at the resolution. 'Co-resolution' referred to discussions in which collaboration may have been the essential component to a resolution. That is, episodes that exhibited deliberation and pooling of linguistic resources; what Donato (1994) labelled 'collective scaffolding', where perhaps neither student could have resolved an episode alone. The hallmark of co-resolution was pair *deliberation* involving suggestion, counter-suggestion, and selection from among options. Excerpt 12 below provides an example of a co-resolved LRE where the resolution was an amalgam of both students' suggestions.

Excerpt 12—Co-resolved (CoR) LRE

- 121 Ken SO, SECOND REASON IS LOSS OF LIFE RHYTHM.
 122 Naoto And maybe 'life rhythm'.
 123 Ken You can't understand?
 124 Naoto From before, I understand your reason. Students don't
 live healthy in summer vacation. But 'life rhythm' hm
 <10> uh I have never heard.
 125 Ken I see. Maybe just 'health'?
 126 Naoto 'Loss of health'? Yes, but but...but you want to talk
 about...connect to how they live in school time and
 vacation. Maybe '*raifusutairu*' (*lifestyle*).
 127 Ken Lifestyle. Good! So, loss of health...no 'lifestyle' <6> or
 'health lifestyle'?
 128 Naoto Hmm, interesting. If <10> uh 'healthy' lifestyle.
 129 Ken Oh, you are clever (both laugh).

Revision: SO, SECOND REASON IS LOSS OF HEALTHY LIFESTYLE.

Deliberation did not necessarily need to result in *agreement* to be coded ‘co-resolved’; as previously seen in Excerpt 7, the ‘author’ of a speech can reject another’s suggestions, but this still is a decision reached after deliberation. Note that extended metatalk (see: Excerpt 10, above) and co-resolution were not always synonymous; Excerpt 13, below provides an example. When Asami was correcting her speech, she produced extended metatalk to provide an explanation-*cum*-justification for her suggested revision (turns 45 and 47). Naho accepts/acknowledges the revision without further contributing to the resolution. As there was no evidence in this LRE of deliberation or co-construction of the revision, it was not coded as co-resolved despite the presence of extended metatalk.

Excerpt 13—Extended (E) metatalk without co-resolution (CoR)

- 44 Asami PEOPLE LIVE IN CITY CAN DISTINGUISH STUDENTS BY ITS
UNIFORM SO...Ah, change ‘its’ by ‘their’, because I’m
saying about students.
- 45 Naho Mm (got it)
- 46 Asami Many student persons. ‘Their uniform’
- 47 Naho Mm (got it)

Revision: [...] DISTINGUISH STUDENTS BY THEIR UNIFORMS.

5.3.5 Inter- and intra-coding of pair-talk data (initial revision)

Figure 5.4 below is an example of an LRE identification-and-analysis tables employed to examine pair-talk during initial revision, and two fully coded LREs.

Figure 5.4 Example LRE identification and analysis table

Original [Line: 4]	THERE ARE A VARIETY OF SLOGAN AND STRONG POINTS FOR EACH SCHOOL.
Initial Revisions	THERE ARE (1) <u>VARIED</u> (2) <u>SLOGANS</u> AND STRONG POINTS FOR EACH SCHOOL.
Pair-talk: <u>Initial</u> revision (M = Momo [author]; C = Chika [revision partner])	

5	M: THERE ARE A VARIETY OF SLOGAN AND STRONG POINTS FOR EACH SCHOOL.	LRE 1			
6	C: Wait. THERE ARE A <u>VARIETY</u> .				
7	M: Hm?				
8	C: 'Variety' is not a correct usage I hear=				
9	M: Really?				
10	C: Maybe it's like XXX XXX (inaudible) like a <i>keiyōshi</i> (<i>adjective</i>).	LRE 2			
11	M: 'There are many kinds'? (suggestion) <i>soka</i> (<i>that's right</i>) <6>				
12	C: I think 'varied'.				
13	M: 'Varied', ah <u>varied</u> ... ok. <15> THERE ARE VARIED SLOGAN, 'slogans'.				
14	C: THERE ARE VARIED SLOGAN...ahhh...'slogans', yes... (writing and reading) THERE ARE VARIED SLOGANS.				
15	M: VARIED SLOGANS AND STRONG POINTS.				
16	C: Good. It's good. FOR EACH SCHOOL, hm (yes, correct).				
LRE Analysis (Initial revision)					
LRE	Focus	Outcome	Quality of Metatalk	Source Initiation	Source Resolution
1	F-LRE (word form)	+/+	E	OI	CoR
2	F-LRE (noun plural)	-/+	L	SI	SR

To ensure reliability of coding, I employed an inter-coding check of the data. The second coder was a native-English speaker colleague holding a doctorate in Applied Linguistics who was familiar with LRE-based studies, but who had limited experience coding LREs himself. Due to the complexity of the data, three rounds of inter-coding were held: (1) an initial training session; (2) inter-coding for LRE focus and outcome; and (3) inter-coding for quality of meta-talk and sources of initiation/resolution. In the initial training session, the second coder was provided with electric copies of (a) a single unmarked transcript of student pair-talk, (b) a scanned color copy of students' revised speech transcript for reference, and (c) a series of corresponding LRE identification/analysis tables with the pairs of unrevised and initially-revised sentences inserted, but otherwise unfilled. An example of (c) is shown in Figure 5.5, below.

Figure 5.5 LRE identification and analysis tables for independent coding

Original [Line 4]	TOO LONG SUMMER VACATION MAKE STUDENTS DANGER PLEASURE, LIKE DRUGS, GAMBLE, AND WOMANIZING.				
Initial revisions [Line 4]	TOO LONG A SUMMER VACATION (1) <u>MAKES</u> STUDENTS (2) <u>TO DO</u> DANGEROUS (3) <u>PLEASURES</u> , LIKE DRUGS, GAMBLE, AND WOMANIZING.				
LREs in relevant pair talk (K = Ken; N = Naoto)					
Analysis of LREs					
LRE	Focus	Outcome	Quality of Metatalk	Source Initiation	Source Resolution
Original [Line 9]	[STUDENTS WANTS TO GO TRIP, BUT IF THEY DON'T HAVE DRIVER LICENCE, THEY WILL UNLICENCE DRIVE.				
Initial revisions [Line 9]	STUDENTS (1) <u>WANT</u> TO GO TRIP, BUT IF THEY DON'T HAVE DRIVER LICENCE, THEY WILL (2) <u>DRIVE CAR WITHOUT LICENCE</u> .				
LREs in relevant pair talk (K = Ken; N = Naoto)					
Analysis of LREs					
LRE	Focus	Outcome	Quality of Metatalk	Source Initiation	Source Resolution

In a computer lab, the second-rater was trained how to paste the relevant portions of pair-talk from transcripts into the tables, how to segment the pair-talk into LREs, and how to code LREs, but only for focus and outcome. The second coder then independently completed the remaining tables. For this training portion of the data, the inter-coder reliability was 88% agreement for the segmentation of LREs, 89% for focus, and 94% for outcome (based on simple percentage agreement). Discrepancies were discussed until 100% agreement was reached.

For the second round of inter-coding, the inter-coder was given 45% of the unmarked transcripts of pair-talk (11 out of 24) with corresponding LRE analysis tables, and scanned color copies of the pair-revised transcripts. The second coder then independently segmented and coded LREs only for foci and outcome. Inter-coder reliability was 94% agreement for segmentation/count of LREs, 92% for focus, and 93% for outcome (using simple percentage agreement). Discrepancies were discussed until 100% agreement was attained for all coding.

For the final round of inter-coding, a training session was held for coding quality of metatalk, source of initiation, source of resolution. The second coder then further independently coded his 45% of the data for these remaining coding categories. Inter-coder reliability was 95% agreement for quality of metatalk, 100% for source of initiation, and 94% for source of resolution (using simple percentage agreement). Discrepancies were discussed until 100% agreement was attained for all coding.

Two weeks later, I did a second ‘fresh’ coding of the pair-talk data from the remaining 12 out of 24 transcripts (i.e., those which had not been inter-rated). The intra-coder reliabilities were (i) 97% for LRE count, (ii) 94% for focus, (iii) 100% for outcome, (iv) 100% for quality of metatalk, (v) 100% for source of initiation, and (vi) 96% for source of resolution.

5.3.6 *Final count of LREs (initial-revision)*

During inter- and intra-coding of LREs, it became apparent that this *raw* count of coded LREs provided a sometimes distorted picture of pair EWL during initial transcript revision because there were cases of pairs revising an *identical* error repeatedly up to four times in total. For example, in Excerpt 12 (p. 78) Naoto and Ken changed LIFE RHYTHM to HEALTHY LIFESTYLE. This initial revision of “life rhythm” had been other-initiated by Naoto, entailed extended metatalk and had been co-resolved. As the pair continued to read through Ken’s transcript, this lexical

item was encountered another three times subsequently, and in each subsequent encounter, was revised to “healthy lifestyle” without extended metatalk. These three subsequent encounters had been (raw) counted as three additional LREs coded as showing limited metatalk; however, this raw count distorted the tally of how often the pair showed signs of extended cognitive engagement with a revision as there was obviously no need for the pair to re-discuss how to revise “life rhythm” when subsequently encountered. Therefore, I re-tallied these four LREs concerned with revising “life rhythm” as a single LRE coded as having been other-initiated and co-resolved after extended metatalk.

I carefully went through all the data for similar cases of multiple LREs that addressed making an identical revision repeatedly. There were no cases of any identical revision being made more than five times or less than three. Revisions considered ‘identical’ were only those made to the *exact same* error/item, not multiple revisions of the same type of error/item. For example, Chika and Momo’s consistent (and required) insertions of the definite article ‘the’ before the word ‘same’ (e.g., WE MUST PAY THE SAME TAXES TO GOVERNMENT or IT IS THE SAME PROBLEM [...]) were counted as an identical revision. Each LRE concerning a separate tokens of an identical error/revision were clustered and re-tallied as a single discontinuous LRE (e.g., LRE 10A, 10B, 10C). If none of the pair’s revision-discussions of the revised item entailed extended metatalk, the clustered LRE was coded as limited, if any of their discussions had entailed extended metatalk, the cluster LRE was coded as extended.

5.4 Analysis and coding of LREs from final-revision pair-talk

After pairs had made their initial revisions, they then turned to the teacher-revised transcripts to make their final revisions. The categories of revisions I (the teacher) had made mirrored those of the students. Dyfluencies were deleted (removed) entirely. Discrete revisions of form and lexis were typed in red font and underlined,

as were any reformulations I had made of a clause/sentence (in this study, teacher-revisions are replicated in **underlined bold** and numbered). Again, I limited myself to making revisions only to forms I felt represent absolute errors in violation of the TL system.

The protocol for final revision was for one student to read the teacher-revised transcript aloud and for the other student to listen and follow along while reading the pair-revised transcript. When differences between the two transcripts were noticed, students were to stop and decide whether further revision of their transcripts was required. In this stage, students made any additional final revision to the transcript exclusively in red pen (replicated in **bold** and numbered) so that revisions derived from teacher feedback could be distinguished from the unmodified, original student revisions. The pair-talk plus any oral contributions from the teacher were transcribed verbatim. The procedures for analysing the interaction data were similar to those above illustrated in section 5.3, but with differences in coding explained below.

5.4.1 Identification of LREs

To identify, segment, and code LREs from final pair-revision, the tables shown above in Figures 5.4 and 5.5 were expanded. These expanded tables first matched the four different versions of the same sentence: (1) the original verbatim unrevised sentence, (2) the sentence as initially revised, (3) the sentence as revised by the teacher, and (4) the final pair-revision of the sentence in light of the teacher feedback. An example is shown below in Figure 5.6.

Figure 5.6 Matched sentences: Initial, teacher, and final revisions

Original [Line 6]	BUT IF WE ONLY HAVE PUBLIC SCHOOL, THEN YOU CAN'T MAKE SURE THEY GET GOOD ENVIRONMENT TO THE STUDENT.
Initial Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) <u>SCHOOLS</u> , THEN YOU CAN'T MAKE SURE-THEY GET GOOD (2) <u>ENVIRONMENTS</u> TO THE STUDENT.

Teacher Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) SCHOOLS , THEN YOU CAN'T MAKE SURE THEY (2) PROVIDE (3) A GOOD ENVIRONMENT (4) FOR (5) EVERY STUDENT.
Final Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) <i>SCHOOLS</i> , THEN YOU CAN'T MAKE SURE THEY (2) PROVIDE GOOD (3) <i>ENVIRONMENTS</i> (4) FOR (5) EVERY STUDENT.

Next, pair-talk where students had discussed the changes the teacher had made to the sentences were copied from the transcripts of student interaction and pasted below into the table. LREs were then identified and matched to revisions.

5.4.2 Coding of LREs

As previously, LREs were defined as any student-interaction where there was an explicit focus on language contained in transcripts of students' speeches. LREs identified from pair-talk during final revision were again coded for (a) focus and prior noticing, (b) outcome and source of revision, (c) quality of metatalk and teacher consultation.

(a) LRE focus and prior noticing

Here LREs were coded for focus as to whether the teacher-revision being discussed focused on grammar form (coded: F-LRE), lexis (L-LRE), or reformulation (REFORM). 'Prior noticing' simply referred to whether the language feature being discussed had been previously noticed and discussed by the pair during the initial revision of the transcript. This category was simply coded 'yes' (Y) or 'no' (N). As some pair discussions during initial revision had led to decisions not to revise a language feature, the segmentation of pair-talk during initial revision into LREs was included in the expanded tables for reference (see: Appendix G).

(b) Outcome and source of revision

These categories were intended to address student uptake of teacher revisions. ‘Outcome’ referred to whether a pair’s decision to adopt a teacher-revision (or not) rendered the final revised version of the transcript accurate or inaccurate. Relative to a pair’s initial revisions, there were three types of teacher revisions:

Type 1: The teacher- and pair-revision of a language feature were identical.

Type 2: Teacher-revision of a language feature that the pair had not discussed as being in (potential) need of revision for accuracy.

Type 3: The teacher- and pair-revision of a language feature were different or the pair had made a revision the teacher had not.

Outcomes, as previously, were coded as either being positive or negative. Students had been instructed to write *single* final revisions and not to write teacher revisions alongside their own (i.e., pairs had to choose), therefore ‘unresolved’ was not a coding category. These codings are shown in Figure 5.7, below.

Figure 5.7 LRE outcome coding for final pair-revision

Outcome	Code	Description
Positive Outcome	(-/ +)	Uptake of teacher-revision renders inaccurate language accurate.
	(+ / +)	Uptake of teacher-revision replaces already accurate language <i>OR</i> decision <u>not</u> to adopt teacher-revision leaves already accurate language unamended.
Negative Outcome	(- / -)	Decision <u>not</u> to adopt teacher-revision leaves inaccurate language unrevised.

Cases of language features revised by the teacher and pair in an identical fashion were often not mentioned by students explicitly, but when such explicit mention was made, these LREs were coded for outcome as (+ / +). *Source of revision* simply referred to whether the final-revision the pair decided upon was an unchanged

revision generated by the pair earlier (i.e., a ‘student-revision’ coded as Stu), or was an adoption of a teacher-revision (i.e., a ‘teacher-revision’ coded as T).

(c) Quality of metatalk and teacher consultation

‘Quality of metatalk’ was, as previously, coded as either limited (L) or extended (E). Here, metatalk was coded as *limited* in cases where:

- Students adopted a teacher’s revision without discussion.
- Students decided not to adopt a teacher’s revision without deliberation (e.g., Student 1: “We don’t need to change” Student 2: “Ok”).
- Students made explicit reference to a language feature merely to note that the student- and teacher-revisions were identical.

Metatalk was coded as *extended* in cases where students:

- Deliberated over whether or not to adopt a teacher’s revision.
- Provided or sought explanation(s) of why/how student- and teacher-revisions differed, or gave reasons why they felt the teacher had made a revision, or the teacher was asked to provide such an explanation.
- Made other meta-comments which clearly indicated an interest or engagement with a teacher’s revision.

An example of an ‘other meta-comment’ is shown below in Excerpt 14: while the LRE showed no signs of deliberation or explanation for the teacher revision, it certainly indicated that Chika was engaged with the teacher-revision and her desire to adopt it.

Excerpt 14—Extended metatalk commentary showing engagement with teacher-revision

- 43 Chika (Reading from pair-revised transcript) AND ALSO uhh THE
PRIVATE SCHOOLS HAVE BETTER ENVIRONMENT THAN
THE PUBLIC SCHOOLS AT THAT POINT. Ok, read,
44 Momo (Reading from teacher-revised transcript) AND ALSO THE

PRIVATE SCHOOLS HAVE A BETTER ENVIRONMENT IN THAT REGARD.

- 45 Chika 'In that regard'. (sighs) (laughs) I want to keep this, his is good but maybe I can't remember it for next time debate. 'In that regard,' 'in that regard, 'in that regard' (repeating for memorization).

Revision: [...] HAVE A BETTER ENVIRONMENT **(1) IN THAT REGARD**.

Teacher consultation merely referred to whether a pair had requested assistance from the teacher to resolve an LRE, and this category was coded simply 'yes' (Y) or 'no' (N).

Figure 5.8 below provides examples of fully coded LREs from pair-talk during final revision of transcripts. (This pair-talk is also included in the example of a full LRE analysis table found in Appendix G). The pair-talk below contained four LREs. LRE 1 (turn 8-9) was triggered by the pair noticing a teacher's revision of lexis (L-LRE). This lexical item had not been previously noticed by the pair and was adopted without discussion. LRE 2 (turn 10-13) was triggered by the pair noticing that their earlier revision of 'environment' into a plural form (F-LRE) differed from the teacher's revision. They deliberated whether to adopt the teacher-revision (*extended metatalk*) and ultimately decided to remain with their initial revision. LREs 3 and 4 involved uptake of teacher revision of language previously unnoticed by the pair. In LRE 3 (turn 17) the teacher's revision of the preposition 'to' → 'for' was adopted without discussion. In LRE 4 (turn 17-19), the meta-talk showed a degree of awareness of the difference between Chika's original use of the determiner 'the' and the teacher's revision changing it to 'every', and therefore this LRE was coded as evidence of extended Metatalk. None of the LREs involved teacher consultation.

Figure 5.8 LRE coding for pair-talk during final transcript revision

Original [Line 6]	BUT IF WE ONLY HAVE PUBLIC SCHOOL, THEN YOU CAN'T MAKE SURE THEY GET GOOD ENVIRONMENT TO THE STUDENT.
Initial Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) <u>SCHOOLS</u> , THEN YOU CAN'T MAKE SURE THEY GET GOOD (2) <u>ENVIRONMENTS</u> TO THE STUDENT.

Teacher Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) <u>SCHOOLS</u> , THEN YOU CAN'T MAKE SURE THEY (2) <u>PROVIDE</u> (3) <u>A GOOD ENVIRONMENT</u> (4) <u>FOR</u> (5) <u>EVERY</u> STUDENT.					
Final Revision [Line 6]	BUT IF WE ONLY HAVE PUBLIC (1) <i>SCHOOLS</i> , THEN YOU CAN'T MAKE SURE THEY (2) PROVIDE GOOD (3) <i>ENVIRONMENTS</i> (4) FOR (5) EVERY STUDENT.					
Pair-talk: <u>Final</u> revision (C = Chika [author]; M = Momo [revision partner])						
11	M: Ok, 6 (line 6) BUT IF WE ONLY HAVE PUBLIC SCHOOLS THEN YOU CAN'T MAKE SURE THEY PROVIDE A GOOD ENVIRONMENT FOR EVERY STUDENT.					
12	C: Why different? <5>					
13	M: Difference is ' <u>PROVIDE A GOOD</u> '. LRE 1					
14	C: Ah (I see). 'Provide'.					
15	M: PROVIDE A GOOD ENVIRONMENT <u>.</u> LRE 2					
16	C: (groans) <6>					
12	M: Do we need change it to 'a good environment' <u>?</u>					
13	C: I think ours is same. But maybe <5> <i>fukasanmeishi, kasanmeishi?</i> (<i>uncountable/countable noun?</i>) I'll check it <40> (consulting dictionary) <i>kasanmeishi (countable noun)</i> so we don't need to change it.					
14	M: Ok, next ENVIRONMENT <u>FOR EVERY</u> STUDENT					
15	C: 'students'?					
16	M: FOR EVERY STUDENT <u></u>					
17	C: 'for' ok...and ' <u>every</u> ' <i>ka! (ah ha)</i> , like, for 'each', each student. LRE 3					
18	M: 'All students'?					
19	C: Yes, but just use it (i.e., 'every'). LRE 3 & 4					
LRE Analysis (<u>Final</u> revision)						
LRE	Focus	Prior Noticing	Outcome	Source Revision	Quality Metatalk	Teacher Consulted
1	L-LRE	N	-/+	T	L	N
2	F-LRE	Y	+/+	Stu	E	N
3	L-LRE	N	-/+	T	L	N
4	F-LRE	N	-/+	T	E	N
Key: L-LRE= Lexical LRE; F-LRE = Form LRE; N=No; Y=Yes; T = Teacher; Stu = student; L = limited; E = extended						

5.4.3 *Inter- and Intra-coding of pair-talk data (final revision)*

To ensure reliability of coding, the same second coder as before was again asked to independently segment and code LREs from pair-talk. After a short training session for the new codings using one of the pair-talk transcripts, the second coder was given 45% of the pair-talk data produced during final revision (i.e., 11 out of 24 copies of unmarked transcripts of pair-talk). Note that these 11 transcripts were of pair-talk from the same learners whose pair-talk during initial revision the inter-rater had coded previously. The second coder was then given the corresponding series of expanded LRE identification and analysis tables (see: Appendix G), and had color copies of the corresponding fully revised speech transcripts for reference. Inter-coder agreement for LRE segmentation and count was 88%, and for:

- a) Focus and prior noticing: 90% and 98% respectively.
- b) Outcome and source of revision: 89% and 100% respectively.
- c) Quality of metatalk and teacher consultation: 94% and 100% respectively.

Discrepancies were discussed until 100% agreement was reached across all coding. A few weeks later, I did a second ‘fresh’ coding of the pair-talk data from the remaining 12 out of 24 transcripts (i.e., those which had not been inter-rated). The intra-coder reliabilities were (i) 97% for LRE count, (ii) 100% for focus and prior noticing, (iii) 100% for outcome and source of revision, (iv) 96% for quality of metatalk, and (v) 100% for teacher consultation.

5.4.4 *Final tally of LREs (final revision)*

Similar to when pairs had made their initial revisions, there were cases of the pairs repeatedly noticing that the teacher had made the same revision to an *identical* error that had been made produced repeatedly. Such cases were each re-tallied as a single discontinuous LRE. As before, if none of the pair’s discussions of a given identical revision entailed extended metatalk, the LRE codes from their first encounter with the item were applied to the entire discontinuous LRE; if any of their

discussions had entailed extended metatalk or co-resolution, the entire discontinuous LRE was coded accordingly.

5.5 Analysis for retention of revisions

In order to investigate how (or whether) the quality of a learner's EWL when making linguistic choices (i.e., revisions) affected language learning-*cum*-consolidation of the language items discussed, it was first necessary to establish whether a revision had been retained. This stage of analysis entailed (a) numbering revisions (b) comparing a student's first (initial/original) delivery of a debate position speech and his/her second delivery (re-delivery) of the same speech.

5.5.1 Counting of revisions

As a first step, any final pair-revisions to 'errors' that, in fact, appeared to have been made to self-transcription mistakes that misrepresented the language the student had actually produced (see Section 5.1) were discounted and removed from the data set. Next, all final pair-revisions to a transcript were renumbered in the order in which they appeared (i.e., the first revision appearing in the transcript was numbered '1', the second '2', and so on). However, identical multiple revisions to the exact same error were numbered, for example, '3A', '3B', '3C', and so on. Therefore, while most revisions were numbered in the order they appeared, there were cases where, for example, 'revision 3A' first appeared in line 4 of the revised transcript and again later as revisions 3B and 3C in lines 11 and 22.

5.5.2 Evidence of retention

The second step was to match the lines from (i) the student's first delivery of their speech, (ii) their initial- and final-revisions to the same lines, and (iii) the corresponding line from their second delivery of a speech. Because all speeches were delivered using the same recommended content-organization (see Appendix C) with the aid of 'debate house' speech notes (see Figure 4.3, p. 54), the ideas (propositional content) contained in both deliveries of a speech were largely the

same and delivered in largely in the same chronological order. A line from a student's first speech delivery was matched to the corresponding line in their second delivery that contained the same propositional content (i.e., the same reason or explanation/example). However, despite similarities in semantic content, there were a number of cases where the same idea from delivery 1 was re-expressed in delivery 2 in such a way that there was no occasion to reproduce a given error or revision. Likewise, there were some ideas contained in delivery 1 that simply were not expressed in delivery 2.

To look for signs of language learning, the line from speech delivery 1 was compared to the corresponding line from delivery 2. Any final revisions to a language item that had been made to delivery 1 that reappeared in delivery 2 were considered as signs of retention-*cum*-learning. For a language item to be considered 'retained/learned' it had to appear exactly as revised. Reformulations were counted by the clause and had to be syntactically the same and error free to be considered 'retained'. Note, however, that dysfluencies—such as, false starts, repetitions, or self-corrections—were not considered 'errors' as these are *not* accuracy-related but rather fluency-related phenomena (Ellis and Barkhuizen, 2005, p. 148). That is, dysfluencies were disregarded and if a revision appeared in the 'final version' of an utterance the revision was considered 'retained'. Evidence of *non*-learning/*non*-retention was taken from cases where the original unrevised form/error reappeared (uncorrected) in the corresponding line of speech delivery 2, or a revision a student attempted to 'reuse' differed in form from that written in the transcript: for example, the revision 'ON THE OTHER HAND' subsequently produced as 'AT THE OTHER HAND' was *not* taken as a sign of retention.

Once a revised item of language from the first delivery of a speech was matched to the item's appearance in the second delivery, the item was then related

to its corresponding LRE. Figures 5.9 and 5.10 illustrate the analysis used in this study for evidence of language retention and learning.

Figure 5.9 Examples of retained revisions

KEN: SCHOOL UNIFORM DEBATE SPEECHES					
Delivery 1 (Line 8)	BUT SOME STUDENTS WHO WANT TO BUY THEIR CLOTHES, THEY WILL NEED BUY SCHOOL UNIFORM PLUS CLOTHES.				
Initial Revisions	BUT SOME STUDENTS WHO WANT TO BUY THEIR CLOTHES, THEY WILL NEED <u>TO</u> BUY SCHOOL <u>UNIFORMS</u> PLUS <u>PRIVATE</u> CLOTHES.				
Teacher Revisions	BUT SOME STUDENTS WHO WANT TO BUY THEIR CLOTHES, THEY WILL NEED TO BUY SCHOOL UNIFORMS PLUS PERSONAL CLOTHES.				
Final Revisions	BUT SOME STUDENTS WHO WANT TO WEAR THEIR CLOTHES, THEY WILL NEED (6) TO BUY SCHOOL (7) UNIFORMS PLUS (8) PERSONAL CLOTHES.				
Delivery 2 (Line 9-10)	SO THE STUDENTS WHO WANT THEIR (8) PERSONAL CLOTHES UH CLOTHES THEY WILL BUY SCHOOL (7) UNIFORMS AND (8) PERSONAL CLOTHES SO THEY WILL PAY THE COST FOR CLOTHES TWO TIMES.				
Retention Results					
A	B	C	D	E	F
Revision number	Focus	Retention	Source Revision	Metatalk	Initiation/ Resolution
6	F-LRE (verb morph)	n/a	n/a	n/a	n/a
7	F-LRE (noun plural)**	O	Stu	(L) + (-)	SI/SR
8	L-LRE (word choice)	O	T	(L) + (E)	SI/TR
Key: n/a = not applicable; ** = Identical student/teacher revision; O = retained; Stu = student; T = teacher; (L)(E) = limited/ extended metatalk; SI = Self- initiated; SR = Student resolved; TR = teacher resolved.					

In the above example, Ken had been arguing the debate position that school uniforms should not be mandatory. The row 'Final Revisions' shows that three final revisions had been made to line 8 of the transcript of the student's (Ken's) initial speech delivery: transcript revisions 6, 7, 8. The row '[speech] Delivery 2', shows that Ken had expressed the same idea to support his position in lines 9-10 of his second delivery. However, due to differences in syntax, in delivery 2 there was no occasion for Ken to employ revision 6 (TO BUY). Therefore, in the 'Retention Results' table at the bottom of the figure, in column C 'Retention' I wrote 'not applicable' (n/a). Ken's redelivery did, however, give rise to occasion to employ revisions 7 and

8 and these revisions did reappear in his second speech. For Revision 7, column B shows that the LRE had concerned a revision for noun plural form (F-LRE) with the double asterisks (**) indicating that the pair- and teacher-revision of the form had been identical, and column C showing that the revision had been retained (O). Column D shows that this final revision was students' and not the teacher's. Column E indicates that it had been a 'single round' revision and was the product of limited (L) metatalk during the pair's initial-revision of the transcript, but was not discussed again (-) during the second final round of revision (coded (L) + (-)). Column F shows that analysis of pair-talk showed this revision's LRE had been self-initiated (SI) and self-resolved (SR).

Analysis for revision 8 showed the following. The final revision/insertion of the adjective **PERSONAL** CLOTHES to clarify Ken's idea did remerge (twice) in his second delivery of the speech and was thus retained (O). Both the pair and the teacher had recognized this need to insert an adjective, but the pair had inserted *PRIVATE* whereas the teacher inserted **PERSONAL**. This difference between the two revisions was noted and discussed during final-revision of transcripts, and therefore this was a 'double round' revision (i.e., discussed both during the making of initial- and final-revisions). The pair's initial-revision had been initiated (and resolved) by Ken with limited metatalk; however, the decision to adopt the teacher's revision was only made after extended discussion. Therefore, this revision was the product of limited plus extended metatalk (L + E) and had been self-initiated but ultimately resolved/revised by the teacher (SI/TR).

Figure 5.10 illustrates revisions which had not been retained. Below, Asami was arguing for the abolition of private schools. Her pair had made three final revisions to line 16 of her first speech delivery: revisions 13, 14, 15. In her re-articulation of her point in speech delivery 2 (lines 19-20), there was no occasion to employ revision 14 (n/a). Revision 13 was not retained (X). This revision was a single

round revision that had been resolved with limited metatalk in round 1 of revision and the teacher's identical revision was not explicitly mention during final revision: metatalk was therefore coded (L + -). This revision had been initiated and resolved by Asami's revision-partner and coded other-initiated/other-resolved (OI/OR). Revision 15 was also not retained (X), and was another single round revision: this time the pair had adopted a teacher revision to a language error not previously noticed. The pair did not deliberate over or discuss the teacher's revision and therefore the metatalk was coded (- + L) and further coded as teacher-initiated/teacher-resolved (TI/TR).

Figure 5.10 Examples of revisions not retained

Figure 5-10 Example of Revisions Not Retained

ASAMI: PRIVATE SCHOOL DEBATE SPEECHES					
Delivery 1 (Line 16)	IF ALL SCHOOLS IS PUBLIC, THEY HAVE SAME EDUCATION.				
Pair Revisions	IF ALL SCHOOLS <u>ARE</u> PUBLIC, THEY HAVE SAME EDUCATION.				
Teacher Revisions	IF ALL SCHOOLS <u>ARE</u> PUBLIC SCHOOLS, THEY <u>PROVIDE THE</u> SAME EDUCATION.				
Final Revisions	IF ALL SCHOOLS (13) <u>ARE</u> PUBLIC SCHOOLS, THEY (14) <u>PROVIDE (15)</u> THE SAME EDUCATION.				
Delivery 2 (Line 19-20)	IF ALL SCHOOLS (13!) IS PUBLIC, IN FACT, ALL STUDENTS GET (15!) SAME EDUCATION.				
Retention Results					
Revision number	Focus	Retention	Source Revision	Metatalk	Noticing
13	F-LRE (verb morph)**	X	Stu	(L) + (-)	OI/OR
14	L-LRE (word choice)	n/a	n/a	n/a	n/a
15	F-LRE (article insert)	X	T	(-) + (L)	TI/TR
Key: ! = revision not retained; n/a = not applicable; ** = Identical student/teacher revision; X = retained; Stu = student; T = teacher; (L) = limited metatalk; OI = Other -initiated; OR = Other resolved; TI = Teacher initiated; TR = teacher resolved.					

Identification of retention of final-revisions was a relatively straightforward process. Data for *double* round revisions (cases where a pair noticed a difference

between their initial-revision and the teacher's revision of the same language feature and had to choose among them), were carefully checked for incidences of a teacher's revision having been adopted but the pair's prior initial-revision having emerged in the second delivery of a speech; and also for cases of a teacher-revision having been rejected but appearing in the student's second delivery. No such cases were found. All retained revisions identified from the data were those the learner had written as their final choice of revision.

5.6 Collection and analysis of interview data

Post-task interviews were conducted to investigate how learners perceived TRTs as language learning activities. Interviews began immediately following completion of the final (3rd) Debate Cycle (see section 4.3) and were held over the course of two weeks. This two-week period corresponded with the university's examination weeks and was the only period in which participants were available for interview before leaving for their two-month vacations. All post-task interviews were transcribed verbatim. These interview transcripts were analyzed using an inductive approach (Strauss and Corbin, 1998). Transcripts were extensively and repeatedly read to extract common themes as to the nature of students' comments. An initial list of the main themes was generated based on the interview questions from the protocols (see Appendix E). However, because not all student comments fit these initial categories (especially those in response to unplanned, spontaneous follow-up interview questions), I expanded and revised the initial categories to better accommodate the data. Once thematic categories/codes were finalized, I reanalysed the interview transcripts using the amended thematic categories. Because interviews had to be held before it was possible to transcribe or analyze any of the learners' pair-interactions during TRTs, it was not possible to ask questions regarding specific learning events; and connections drawn between learners' perceptions of TRTs and their TRT performances were made *post-facto*.

6 CHAPTER 6: ENGAGEMENT WITH LANGUAGE DURING TRANSCRIPT REVISION

This chapter reports the results of the study by answering the first research question: What was the quality of learner engagement with language (EWL) during transcript revision? Students' verbalization processes during transcript revision tasks (TRTs) were explored by copying and pasting the coded data from the study's LRE identification/analysis tables into Microsoft Excel and using the "sort data" feature to: (a) look for whole-class trends in EWL common, and (b) look for any differences in engagement between pairs.

6.1 Results for Research Question 1a: Whole class trends during initial-revision

Here whole class trends are reported regarding the foci, outcomes and nature of the LREs produced by pairs during initial-revision of transcripts. Across three TRTs (two transcripts revised per TRT), the four pairs produced a total of 448 LREs. Of this total, 192 LREs were identified as concerning deletions of dysfluencies (on average eight per revision of a transcript) and six as concerned with addition of new content. These LREs were not analysed in the study, leaving 290 LREs remaining which addressed making repairs or improvements to language already contained in the transcripts of learner-oral performance.

6.1.1 Number of revision-related LREs

Table 6.1 below summarizes the number of revision related LREs produced by the four pairs across each of the three TRTs performed. In the table, each Pair/TRT is further subdivided to show the number of LREs produced when a given student's transcript was being revised by the pair. For example, the table shows that during TRT 1, Pair 1 generated 7 LREs when revising Naoto's transcript and 16 LREs when revising Ken's transcript.

Table 6.1 Number of revision-related LREs across TRTs

Pair / Student		TRT 1 Number of LREs	TRT 2 Number of LREs	TRT 3 Number of LREs	Total LREs
1	Naoto	7	10	10	27
	Ken	16	16	14	46
2	Chika	9	15	14	38
	Momo	17	11	15	43
3	Asami	10	6	14	25
	Naho	13	4	15	32
4	Aki	23	13	19	55
	Yuta	6	10	8	24
Grand Totals		98	88	109	290

The above shows that, when revising transcripts of on average 400 words in length of speeches that on average were 7 minutes in length, the largest number of revisions discussed was 23 by Pair 4 (Aki and Yuta) during TRT 1. This same pair, across three 20 minute TRT sessions revising Yuta's transcripts produced only 24 LREs *in total*. The mean average for the class was 12 revision related LREs per transcript. The number of LREs pairs produced did not increase with time. A comparable number of revisions and LREs produced were produced at the start of the study in TRT 1 as were in TRT 3 by the end of the study. Although previous TRT studies were either of a single TRT or did not report learners' noticing/revision rates over time, by way of comparison, the pairs in Lynch's (2001) study on average produced (and presumably discussed) 28 revisions to transcripts of 2 minute oral

performances; the triad of Japanese learners in Mennim's (2003) study noticed 49 possible errors in transcripts of 5 minutes of speech; and the Japanese learner pairs in Stillwell *et al.* (2010) on average made 15 revisions to transcripts of 3 minutes of oral performance. Compared to prior TRT studies, my learners on average produced fewer revisions/LREs.

6.1.2 *Foci of LREs*

Table 6.2 shows that pairs produced form-based LREs most frequently (68.62%), followed by lexis (22.14%), and reformulations of sentences or clauses (8.96%). As in prior TRT studies, this study found intermediate-level university learners of English consistently focused attention more often on grammatical items rather than on lexis. However, learners in this study had a higher proportion of LREs focused on lexis (22.14%) compared to the learners in Lynch (2001) and Stillwell *et al.* (2010) where (respectively) only 7% and less than 1% of learner revisions were lexical.

Table 6.2 Summary of foci of LREs (initial transcript revision)

<i>Type</i>	<i>Number</i>	<i>% Total LREs</i>
Form-LREs	199	68.62%
Lexical-LREs	65	22.14%
Reformulations	26	8.96%
<i>Total LREs</i>	<i>290</i>	<i>100.00%</i>

Table 6.3 shows the specific foci of the form- and lexical- LREs. Revision of verb morphology or auxiliary verbs comprised 43.7% of all form-LREs with issues of subject-verb agreement comprising just over a quarter of all F-LREs; followed by noun pluralisation (16.08%) and conjunctions (11%). Closer examination of LREs concerning conjunctions revealed that these revisions were seldom corrections of errors (e.g., using 'so' when the argument being made required 'but'), but rather stylistic changes to avoid overuse of the same conjunction (e.g., changing 'and' to 'plus' or 'in addition'). As for lexical-LREs, just over a third (36.92%) dealt with choice or insertion of prepositions with the remainder addressing word choice/meaning.

Table 6.3 Foci of Form-LREs and Lexical-LREs (initial transcript revision)

Form-LREs	No.	% of Total F-LREs
Noun Pluralisation	32	16.08%
Verbs		
· <i>Morphology (Subject-verb agreement)</i>	55	27.63%
· <i>Morphology (Tense/aspect)</i>	22	11.05%
· <i>Change insert auxiliary verb</i>	10	5.02 %
Conjunctions/Conjunctive adverbs	22	11.05%
Determiners		
· <i>Articles</i>	15	7.53%
· <i>Other</i>	6	3.01%
Word form (e.g., 'equal' vs. 'equality')	19	9.54%
Other		
· <i>Pronouns</i>	6	3.01%
· <i>Negative construction</i>	4	2.01%
· <i>Genitive 's'</i>	3	1.5%
· <i>Adjectives</i>	2	1%
· <i>Insert missing subject or object</i>	2	1%
· <i>Choice of Conditionals</i>	1	0.5%
Total F-LREs	199	100%
Lexical-LREs	No.	% Total L-LREs
Prepositions	24	36.92%
Word choice/meaning	41	63.07%
Total L-LREs	65	100.00%

6.1.3 LRE outcome

Table 6.4 shows that the vast majority of LREs were resolved (only 2.06% were unresolved) and had positive outcomes regardless of focus. Form-LREs had positive outcomes 92.96% of the time; 80% of lexical-LREs had positive outcomes, as did reformulations (80.76%). Only 8.96% of revision resolutions had negative outcomes and only 1.37% rendered the language contained in transcripts less accurate (+/-) than before. However, only 76.89% of form- and 40% of lexically-focused LREs actually revised an inaccurate language feature (-/+); the remaining F-LREs and L-LREs lead to accurate modification to something already correct (+/+). Similarly, not

all reformulations had been ‘necessary’ and 26.92% of these revisions had (+/+) outcomes.

Table 6.4 LRE outcomes by focus (initial transcript revision)

Focus of LRE	Positive outcomes		Negative outcomes		Unresolved (?)	Total LREs
	(-/+) (76.89%)	(+/+) (16.08%)	(-/-) (6.03%)	(+/-) (0.50%)		
Form-LREs (% of F-LREs)	153 (76.89%)	32 (16.08%)	12 (6.03%)	1 (0.50%)	1 (0.50%)	199
Lexical-LREs (% of L-LREs)	26 (40.00%)	26 (40.00%)	7 (10.76%)	3 (4.61%)	3 (4.61%)	65
Reformulations (% of Reformulation)	14 (53.84%)	7 (26.92%)	3 (11.53%)	0 (0.00%)	2 (7.69%)	26
<i>Total (% of total LREs)</i>	<i>193 (66.55%)</i>	<i>65 (22.41%)</i>	<i>22 (7.85%)</i>	<i>4 (1.37%)</i>	<i>6 (2.06%)</i>	<i>290 (100%)</i>

6.1.4 Quality of metatalk and LRE outcomes

In this study, quality of metatalk was used as a proxy for cognitive EWL; that is, as a gauge of the cognitive effort put into revising a language. Table 6.5 demonstrates that 65.17% (189/290) of LREs had involved limited metatalk, and of these 95.23% (180/189) had positive outcomes. In other words, 62% (180/290) of all language issues could be resolved accurately with only limited metatalk. In contrast, only 34.82% (101/290) of LREs had involved extended metatalk, of which 77.22% (78/101) had positive outcomes.

Table 6.5 Quality of metatalk and LRE outcomes (initial transcript revision)

Quality of LRE metatalk	Positive outcomes		Negative outcomes		Unresolved (?)	Total LREs
	(-/+)	(+/+)	(-/-)	(+/-)		
Extended metatalk (% of Extended)	47 (46.53%)	31 (30.69%)	14 (13.86%)	3 (2.97%)	6 (5.94%)	101
Limited metatalk (% of Limited)	146 (77.24%)	34 (17.89%)	8 (4.23%)	1 (0.5%)	0 (0.00%)	189
Total (% total LREs)	193 (66.55%)	65 (22.41%)	22 (7.85%)	4 (1.37%)	6 (2.06%)	290 (100%)

Results presented in Table 6.6, below show that form-LREs appeared easier for the learners to resolve insofar as only 30.15% of all F-LREs entailed extended metatalk, and the remaining 69.84% (139/199) F-LREs resolved with limited metatalk. The exceptions were revisions of conjunctions/conjunctive adverbs of which 63.63% (14/22) entailed extended metatalk, and changes to word form where 57.89% (11/19) also entailed extended metatalk. However, closer examination of transcribed metatalk revealed that extended metatalk during revision of conjunctions most often was merely a learner explaining that they were proposing the revision for stylistic reasons (i.e., for the sake of lexical variety) and not because the extant conjunction was inaccurate. Resolving lexical-LREs, especially discussion over word choice and meaning, more often entailed extended metatalk: 46.15% (30/65) of lexical revisions were the product of extended metatalk. Reformulations most often entailed extended discussion (18/26 or 69.23%); unsurprisingly given that reformulation usually entailed discussion of a learner's intended idea/meaning and joint re-composition to make it more comprehensible.

Table 6.6 Quality of metatalk and LRE focus during initial transcript revision

Form-LREs	No.	Limited No. (% of No.)	Extended No. (% of No.)
Noun Pluralisation	32	28 (87.50)	4 (12.50%)
Verbs			
· <i>Morphology (Subject-verb agreement)</i>	55	46 (83.63%)	9 (16.36%)
· <i>Morphology (Tense/aspect)</i>	22	16 (72.72%)	6 (27.27%)
· <i>Change or insert auxiliary verb</i>	10	7 (70.00%)	3 (30.00%)
Conjunctions/Conjunctive adverbs	22	8 (36.36%)	14 (63.63%)
Determiners			
· <i>Articles</i>	15	13 (86.66%)	2 (13.33%)
· <i>Other</i>	6	1 (16.66%)	5 (83.33%)
Word form (e.g., 'equal' vs. 'equality')	19	8 (42.10%)	11 (57.89%)
Other			
· <i>Pronouns</i>	6	6 (100.00%)	0 (0.00%)
· <i>Negative construction</i>	4	1 (25.00%)	3 (75.00%)
· <i>Genitive 's'</i>	3	2 (66.66%)	1 (33.33%)
· <i>Adjective insert</i>	2	0 (0.00%)	2 (100.00%)
· <i>Insert missing subject or object</i>	2	0 (0.00%)	2 (100.00%)
· <i>Choice of Conditionals</i>	1	0 (0.00%)	1 (100.00%)
Total F-LREs	199	139 (69.84%)	60 (30.15%)
Lexical-LREs	No.	Limited No. (% of No.)	Extended No. (% of No.)
Prepositions	24	17 (70.83%)	7 (29.16%)
Word choice/meaning	41	18 (43.90%)	23 (56.09%)
Total L-LREs	65	35 (53.84%)	30 (46.15%)
Reformulations	No.	Limited No. (% of No.)	Extended No. (% of No.)
Reformulations	26	8 (30.76%)	18 (69.23%)
Total Reformulations	26	8 (30.76%)	18 (69.23%)

6.1.5 Joint participation in LRE resolution

Another window into the nature of learners' EWL during initial pair-revision was provided by looking at the sources of LRE initiation and resolution. These codings

allowed for a general overview of the extent of a learner's participation in revision of their own transcripts relative to their participation in the revision of their partner's transcripts. Table 6.7 below groups LREs into two categories. The first, 'Individual initiation and resolution', refers to episodes where both LRE initiation and resolution was entirely the work of either the student whose transcript was being revised (Self-initiated/Self-resolved) or their revision partner (Other-initiated/Other-resolved). The second category, 'Co-resolution or Mutual metatalk' refers to episodes which were either co-resolved or those where one participant initiated the LRE and the other resolved it. As Table 6.7 shows, most LREs were not the product of joint resolution. On average 69.31 % (201/290) of all LREs were entirely the products of an individual participant: 47.58% of all LREs were self-initiated and resolved, and 20.83% were other-initiated and resolved. Only 30.68% (89/290) of episodes entailed a degree of joint participation, and on average only 21.72% of all LREs outcomes were the product of co-resolution (including the six unresolved LREs).

Table 6.7 Individual vs. joint initiation/resolution of LREs

Total LRES	Individual initiation and resolution		Co-resolution or Mutual metatalk	
	Self-initiated / Self-resolved (% total LREs)	Other-initiated/ Other-resolved (% total LREs)	Co- Resolved (% total LREs)	Mutual limited metatalk (% total LREs)
290	138/290 (47.58%)	63/290 (21.72%)	*63/290 (21.72%) * 6 unresolved	26/290 (8.96%)

6.2 Results for Research Question 1a: Pair-by-pair analysis

Although analysis of LRE data revealed broad class-level trends in EWL, when examined on a pair-by-pair basis the data uncovered notable differences between learners in how they performed TRTs to create opportunities to engage with language. To begin accounting for these differences required examination of learner

pair-talk beyond analysis of coded LRE data. Therefore, in addition to a pair-by-pair examination of coded data, a closer more descriptive analysis of learner pair-talk was also undertaken.

6.2.1 *Pair 1: Naoto and Ken*

In Pair 1 there was a significant difference in TL proficiency between participants. Naoto was an older student (21 years old compared to the usual 18/19 year old freshmen student) who had 9 years of English study—three of which were at an English medium international school from ages 7-10—and Naoto’s TOEFL IPT English proficiency score was 520. Ken (18 years old) had a lower proficiency test score of 460 and a total of six years of English study in Japanese middle and high school.

Table 6.8 below, tallies the pair’s LREs in terms of their sources of initiation and resolution. Given the differences in age, experience, and TL proficiency, it was unsurprising that the data revealed that, across the three TRTs performed, Naoto had initiated and resolved 46% of all revision episodes: 70% (19/27) of all language episodes concerning revision of his transcripts, and 30% (14/46) of those concerning revision of Ken’s transcripts. However, Ken initiated and made 32% (15/46) of all revisions to his transcripts and 23% (17/73)—nearly a quarter—of all revisions were co-resolved (including six unresolved attempts at co-resolution). These results indicate that Naoto did not assume what Storch (2002) would describe as a dominant role in the revision process, nor Ken assume a passive one. Rather, the data suggested that, overall, Naoto and Ken had taken (respectively) expert and novice roles during transcript revision.

Table 6.8 Pair 1: Individual vs. joint initiation/resolution of LREs

Student Pair		Total LRES	Individual initiation and resolution		Co-resolution or Mutual metatalk	
			SI/SR	OI/OR	Co- Resolved	Mutual metatalk
1	Naoto	27	19/27 (70.37%)	1/27 (3.70%)	7/27 (25.92%)	0 (0.00%)

	Ken	46	15/46 (32.60%)	14/46 (30.43%)	*10/46 (21.73%) *1 unresolved	7/46 (15.21%)
Key: SI/SR = self-initiated/self-resolved; OI/OR = other-initiated/other-resolved						

However, the roles assumed by Naoto and Ken were affected by whose transcripts the pair was revising. Pair 1 began all TRTs with revision of Ken's transcript(s) followed by revision of Naoto's. Interestingly, it was Ken who had been recorded as establishing this order of transcript revision at the start of TRT 1. Ken further recommended that, in his words, "a smart way" of revising, was to read each line of his transcript twice: first he would read aloud and make changes, and then Naoto re-read the (now revised) line and make additional changes. Excerpt 15 below from TRT 1 illustrates their process.

Excerpt 15

- 69 Ken MY THIRD REASON NEGLECT STUDIES, oh ok, 'is' neglect studies, studies...wait. Please.
- 70 Naoto Ok.
- 71 Ken Neglect 'of' studies.
- 72 Naoto Better
- 73 Ken So let's change it.
- 74 Naoto I say yes (while writing revision) 'Neglect of studies'.
- 75 Ken Ok, more?
- 76 Naoto MY THIRD REASON IS NEGLECT OF STUDIES. No.

If one criteria for identifying whether one participant is acting as a pair's 'expert' is determining initiates most discussions (Storch, 2002), in the excerpt above Ken is clearly acting as the expert. This process proposed by Ken guaranteed, in effect, that there would be a number of opportunities (such as those in the excerpt above) for Ken to have the autonomy to lead and initiate revisions. Also, the above

excerpt—and Naoto's adherence to Ken's suggested revision process—illustrate Naoto's willingness to take direction from his partner, a sign that there was considerable social equality between the two learners.

It was not unusual for the role of 'expert' to alternate during revision of Ken's transcripts. Ken initiated a significant one third of form-LRE revisions, but very rarely initiated revisions of lexis which were most often Naoto's purview when revising, as shown in Excerpt 16 below from TRT 3. In the excerpt, Ken raised a rather sophisticated grammar point (subjunctive mood) in Turn 117. Naoto appeared not to have noticed this aspect of Ken's speech, and appeared rather excited when Ken pointed it out, finding Ken's revision to be 'cool' (Turn 118). With only a modicum of assistance from Naoto (Turn 122), this LRE was co-resolved accurately. Upon re-reading the revised portion of the transcript, Naoto (Turn 126) acted as the 'lexical expert' and questioned Ken's use of words 'riches and poors'. He explains the problem (Turn 128) and suggests a revision (Turn 130) which Ken gratefully accepts.

Excerpt 16

- 117 Ken AND THERE IS NOT GAP BETWEEN RICHES AND POORS.
Wait, wait, wait. Ah, line, on my line, this is a,
ah...*kateihou* (*subjunctive mood*), so uh, 'there would not
be gap'
- 118 Naoto Ah, Ah, Ah. Yes! I follow. *Kakkoi* (*cool/handsome/nifty*)
- 119 Ken I think its better. Here is 'if' (in the previous line at start of
Ken's argument).
- 120 Naoto 'wouldn't be gap'.
- 121 Ken 'would not be gap' <7>
- 122 Naoto So, I will (read this) AND THERE WOULD NOT BE GAP 'gap'
uh.. 'a'(!) gap.
- 123 Ken *So ja, dake* (*oh right*) 'a gap'.
- 123 Naoto Yes, I think so.
- 125 Ken Ah yes.
- 126 Naoto BETWEEN RICHES AND POORS. 'Riches' means group of

rich? And group of poor means 'poors'?

127 Ken Hm, yes.

128 Naoto I think 'riches' means money, not people.

129 Ken Oh really? (laughs)

130 Naoto I think so. Ah, 'rich people and poor people'.

131 Ken Hm (yes) I see, thank you. 'Rich people and poor people'
<8>, Oh, I wanted to use 'poors' XXX XXX (inaudible).

132 Naoto (Laughs)

133 Ken I am from a family of poors (laughs)

134 Naoto (Laughs) My hometown is (laughs) nobody are riches, a
community of poors. (Both laugh).

The excerpt above illustrates the often exhibited willingness of this pair to engage with each other's ideas and to provide feedback and explanations. Turns 131-134 illustrate another recurring aspect of this pair's interactions: these students often found humor in their errors. This ability to joke about mistakes likely contributed to this pair's ability to sustain their interactions and engagement with the language contained in the transcripts.

When revising Naoto's transcripts, Naoto did overwhelmingly control the task in that he initiated 85% (23/27) of revisions and self-resolved 70% (19/27) of them. In part, I ascribe this to the pair continuing to follow the same revision process as when revising Ken's transcripts, but in reverse. That is, Naoto would read and revise a line/portion of his transcript, and then ask Ken to re-read and suggest additional revisions. For the most part, Ken simply could not identify any features for revision that Naoto had not already revised, and only identified four such features over three TRTs (three in TRT 2, and one in TRT 3). Although Ken's ability to identify language problems Naoto had missed was limited, whenever Naoto was unsure of a revision, he invariably solicited Ken's opinion, as this excerpt below from TRT 3:

Excerpt 17

- 90 Naoto IF WE HAVE 100 CLASSES 200 000 YEN, IT IS BETTER THAN 50 CLASSES 200 000 YEN. Do you think 'for' 200 000 yen? I'm not certain but I don't know whether it is wrong or not.
- 91 Ken Maybe 'in', 100 classes in 200 000 yen.
- 92 Naoto For 200 000, in 200 000, for 200, 000 <6> Maybe 'in' is better but, uh, I don't know. You choose.
- 93 Ken Me? Maybe 'in'?
- 94 Naoto So please change into 'in'.

Furthermore, even when not questioning the revisions Naoto made to his own transcript nor directly contributing to resolutions, Ken exhibited EWL by often articulating the reasons behind Naoto's revisions, as in Excerpt 18 below from TRT 3.

Excerpt 18

- 37 Naoto I WANT TO MENTION THAT THERE IS A DIFFERENT NATURE BETWEEN THESE CHOICES. Hmm (thinking) 'Different nature' change to 'difference'.
- 38 Ken A 'difference nature'.
- 39 Naoto 'Difference nature' *ja nai te (not that)*, just I want 'difference'. No 'nature', cut that, only 'difference'.
- 40 Ken Ok...(writes and reads revision) A DIFFERENCE BETWEEN THESE CHOICES, oh, I see your mean, there are two choices, not the same uh...because different choice.
- 41 Naoto Hm (yes). Difference is important in my idea (i.e., debate point).

Out of the 27 LREs produced when revising Naoto's transcripts, 15 involved extended metatalk of which 8 did so because Ken had articulated his understanding of Naoto's revisions. Table 6.9 shows that this pair's ratio of LREs involving limited

versus extended metatalk was higher than the class average: 41% (30/73) of their LREs had involved extended metatalk compared to the class average of 35%.

Table 6.9 Pair 1: Quality of LRE metatalk across TRTs

Student Pair		TRT 1		TRT 2		TRT 3		No. / Total LREs (% of Total LREs)	
		E	L	E	L	E	L	Extended	Limited
1	Naoto	3	4	4	6	5	5	12/27 (44%)	15/27 (56%)
	Ken	7	9	4	12	7	7	18/46 (39%)	28/46 (60%)
Grand pair total across TRTs								30/73 (41%)	43/73 (58%)
Key: E = Extended metatalk; L = limited metatalk									

In sum, while Naoto had led the revision of transcripts overall, the pair's interactions were not one-sided, with Ken creating his own opportunities to engage with language. The learners were socially engaged and, affectively speaking, appeared to have enjoyed transcript revision. Their apparent EWL, plus their agreed upon process of reading every line twice, led this pair to spend the full 20 minutes allotted for revising each transcript in all three TRTs.

6.2.2 Pair 2: Chika and Momo

Pair 2 was another mixed TL proficiency pair. Chika had a TOEFL ITP score of 500 versus Momo's lower 475. Chika had also spent 6 months in America as an exchange student at age 17. There was no evidence of this pair having settled upon a fixed order in which to revise transcripts (e.g., TRT 1 began with reading of Chika's transcript, TRTs 2 and 3 with Momo's). In all TRTs, the student who had delivered the speech read the transcript aloud and initiated LREs as language features were self-

noticed, with the partner also jumping in to initiate LREs whenever a language feature was noticed by the partner.

However, across the three TRTs it was Chika who initiated most of the noticing. As shown in Table 6.10 below, Chika had initiated 76% of all LREs during revision of her own transcript, and initiated 63% of LREs during revision of Momo's transcripts.

Table 6.10 Pair 2: Initiation of LREs across three TRTs

Student Pair	TRT 1		TRT 2		TRT 3		No. / Total LREs (% of Total)	
	Self	Other	Self	Other	Self	Other	Self	Other
Chika	8	1	14	2	7	6	29/38 (76%)	9/38 (24%)
Momo	7	10	4	7	5	10	16/43 (37%)	27/43 (63%)

There were clear indications that Chika had noticed and remembered a number of language features from self-transcription that she wanted to revise. For example, reading of her transcript in TRT 1 commenced with Chika saying, "I have to say it was a horrible, horrible speech debate. There are many, I want to change many things. But if you find, like, problem, just tell me". In each session when reading from her transcript Chika read very quickly and often made revisions in rapid succession; and Momo often had some difficulty keeping pace just writing Chika's revisions with little time to notice any remaining points for change. An example is shown in Excerpt 19 below from TRT 2:

Excerpt 19

- 38 Chika AND FOR NEEDY, 'the' needy, THE NEEDY STUDENTS,
ah...change to 'buying' CLOTHES IS GOING TO BE
FINANCIAL BURDEN, 'a financial burden'.
- 40 Momo One more?
- 41 Chika Ok. Line 8. FOR THE ok? NEEDY STUDENTS BUYING=

- 42 Momo 'buying'?
- 42 Chika Mm (yes) BUYING CLOTHES IS GOING TO BE A FINANCIAL BURDEN.
- 43 Momo 'a financial burden', ok.
- 44 Chika Next (continues reading)

Chika would also often jump in with a suggested revision when Momo was reading her transcripts aloud before Momo had a chance to complete reading a given line of transcript, as shown in Excerpt 20 below from TRT 1:

Excerpt 20

- 21 Momo SO THE TIME OF MAKE MONEY IS=
- 22 Chika 'making' money, maybe, I think.
- 23 Momo Eh? uh, but I tried use 'make' but...stop, stop.
- 24 Chika Ah, so the time of make money is important...time to make money, students need, right?
- 25 Momo Yes, yes, yes.
- 26 Chika Ok...so I think 'time making money' is, um IMPORTANT FOR STUDENTS.
- 27 Momo [no response from Momo but sound of pencil writing revision into her transcript] Ok, let's (keep reading).

In all, Chika had self- initiated and resolved 52.63% of revisions made to her transcripts, and had initiated and resolved 46.51% of revisions made to Momo's transcripts, as shown in Table 6.11.

Table 6.11 Pair 2: Individual vs. joint initiation/resolution of LREs

Student Pair		Total LRES	Individual initiation and resolution		Co-resolution or Mutual metatalk	
			SI/SR	OI/OR	Co- Resolved	Mutual metatalk
2	Chika	38	20 (52.63%)	5 (13.15%)	*8 (21.05%) *2 unresolved	5 (13.15%)

	Momo	43	8 (18.60%)	20 (46.51%)	*10 (23.25%) *2 unresolved	5 (11.62%)
Key: SI/SR = self-initiated/self-resolved; OI/OR = other-initiated/other-resolved						

However, the results presented in Table 6.11 above should not be taken to indicate that Chika had entirely dominated the pair interaction. Firstly, 22% of all of their LREs were co-resolved (see table above). In fact, in every TRT, Chika and Momo spent the full 20 minutes allotted to revise a transcript reading through each transcript twice. While during the first read through of a transcript Chika did overwhelmingly lead the revision process, in the second reading Momo became able to spot language points that Chika had missed. More importantly, in the second reading of a transcript, Momo questioned a number a number of Chika's proposed revisions, which meant the pair produced discontinuous LREs that had been initially 'resolved' by Chika but ultimately were co-resolved and amended. An example is shown below in Excerpt 21 from TRT 3 (Chika's original revision is shown below in *HANDWRITING* font):

Excerpt 21

- 156 Chika *ON THE OTHER HAND* EVEN IF WE MAKE ALL
SCHOOLS BE PUBLIC, what?
- 157 Momo Mm, I think 'on other hand' is wrong.
- 158 Chika Eh? Why?
- 159 Momo *Eto (um)* <7> this is another reason for private schools.
- 160 Chika Yes, because will still be rich students and poor students.
- 161 Momo But, ah this (on the other hand) means 'but', mm you
changed this 'but'. This is not 'but' it is another reason.
- 162 Chika Ah! So, like, it is not uh logically?
- 163 Momo Maybe. Maybe this is 'and'.
- 164 Chika Yes, so <7> you are right. But (laughs) 'and' is...I don't like
'and'.
- 165 Momo 'Plus'?

- 166 Chika Or, like 'additional'?
- 167 Momo 'In'?
- 168 Chika 'In additional'.
- 169 Momo Maybe.

Additionally, Table 6.12 below shows that this pair, like Naoto and Ken (Pair 1), had a higher than class average proportion of LREs exhibiting extended versus limited metatalk. 46% (37/81) of Chika and Momo's LREs exhibited extended metatalk compared to the class average 35%. Cases of extended metatalk that were not part of LRE co-resolution *per se*, may have represented cases of a student using language to mediate their own thought by thinking aloud when recommending a revision, or an implicit seeking of affirmation for their revision (although these 'explanations/justifications were not said with a questioning tone). Additionally, these learners would also give unsolicited 'explanations' of revisions the partner was making to her own transcript. These may have been made because the learner had not have fully understood the reason for their partner's revision and was seeking clarification; however, such explanations were also not said in a questioning tone. These explanations may simply have been given by learners to signal engagement in revision of their partner's transcripts rather than as contributions to LRE resolution *per se*. In either case, Chika and Momo did seem equally engaged with the process of both revising their own and their partner's transcripts.

Table 6.12 Pair 2: Quality of LRE metatalk across TRTs

Student Pair		TRT 1		TRT 2		TRT 3		No. / Total LREs (% of Total LREs)	
		E	L	E	L	E	L	Extended	Limited
2	Chika	3	6	8	7	5	9	16/38 (42%)	22/38 (58%)
	Momo	9	8	5	6	7	8	21/43 (49%)	22/43 (51%)

Grand pair total across TRTs	37/81 (46%)	44/81 (54%)
Key: E = Extended metatalk; L = limited metatalk		

In addition to appearing to have been cognitively engaged with revision, this pair's social engagement appeared, much like with Pair 1, to have been maintained, much like with Pair 1, by 'affective scaffolding' (Villamil and Guerrero, 1996) such as laughing and gently ribbing each other over mistakes, as shown in Excerpt 22 below from TRT 3:

Excerpt 22

- 79 Chika Arrg! Same mistake before, 'be disappeared'. Why do I always say this thing?
- 80 Momo Stupid. (both laugh)
- 81 Chika It's, (laughs) you know, hard to remember when saying speech.
- 82 Momo *Gambatte (good luck/you can do it)* stupid (both laugh).

6.2.3 Pair 3: Naho and Asami

Unlike the pairs discussed so far, Naho and Asami had much closer English proficiency scores: Naho with TOEFL ITP scores of 466 and Asami 458, with both having had the standard 6 years of English study at Japanese middle and high school.

In Table 6.13 below, the LRE analysis showed an overall pattern of LRE initiation similar to those of Pairs 1 and 2; that is one participant—Naho—initiated the majority of LREs (68% or 39/57). The percentage of co-resolved LREs for Naho and Asami were comparable to those of Pairs 1 and 2: this pair had co-resolved or attempted to co-resolve 21% (12/57) of LREs compared to 23% and 22% for Pairs 1 and 2 respectively.

Table 6.13 Pair 3: Individual vs. joint initiation/resolution of LREs

Student Pair		Total LRES	Individual initiation and resolution		Joint initiation / resolution	
			SI/SR	OI/OR	Co- Resolved	Mutual metatalk
3	Naho	32	24 (68.75%)	2 (6.25%)	5* (15.62%) *1 unresolved	1 (3.12%)
	Asami	25	6 (24.00%)	10 (40.00%)	7 (28.00%)	2 (8.00%)
Key: SI/SR = self-initiated/self-resolved; OI/OR = other-initiated/other-resolved						

Here, however, the similarities between this pair and the previous two pairs ended. Firstly, Naho appeared to dominate the task to a larger degree than did the ‘leaders’ in the previous pairs. Naho had initiated *and* resolved the 60% (34/57) of all the LREs produced by the pair, compared to Naoto having had having had initiated and resolved only 47% all of Pair 1’s LREs and Chika having done so for 49% of the LREs in Pair 2.

Table 6.14 below shows that the frequency with which Naho and Asami produced extended metatalk was also comparatively low: only 28% of the LREs produced by Pair 3 had involved extended metatalk, compared to the 41% and 46% of LREs from Pairs 1 and 2 respectively. While Asami and Naho’s more perfunctory pair-talk indicated a lesser degree of EWL than that demonstrated by their peers in Pairs 1 and 2, this comparatively low proportion of extended LREs may only indicate that the majority of language features noticed were easy for Naho and Asami to revise and did not require elaboration or deliberation.

Table 6.14 Pair 3: Quality of LRE metatalk across TRTs

Student Pair		TRT 1		TRT 2		TRT 3		No. / Total LREs (% of Total LREs)	
		E	L	E	L	E	L	Extended	Limited
3	Naho	4	9	1	3	4	11	9/32 (28%)	23/32 (72%)
	Asami	1	9	3	5	3	4	7/25 (28%)	18/25 (72%)
Grand pair total across TRTs								16/57 (28%)	41/57 (72%)
Key: E = Extended metatalk; L = limited metatalk									

This pair's mostly perfunctory pair-talk seemed to indicate a low degree of social engagement as it signaled limited effort to maintain interaction, which involves not only the sustaining of interaction but also caring for its quality (Svalberg, 2009, p. 246). Signs of affective engagement with the revision process—such as laughter or excitement—were notably absent from Asami and Naho's pair-talk. An overall indicator of a lower degree of EWL was that these learners spent only roughly half the amount of time their peers took reading and revising transcripts of comparable length to those that had been produced by their peers. Unlike Pairs 1 and 2, Asami and Naho read through their transcripts only once, and spent only approximately nine to 13 minutes revising each transcript.

6.2.4 Pair 4: Aki and Yuta

Aki and Yuta had similar TOEFL ITP proficiency scores of 480 and 473 respectively. Unlike the pairs previously discussed, Pair 4 was the only pair where the nature of the learners' EWL changed over the course of the study. As Table 6.15 illustrates, in TRT 1 when revising Aki's transcript 30% (7/23) of LREs were co-resolved, and 66% (4/6) of those produced during revision of Yuta's transcripts; however, in TRT 3 *zero* LREs produced by the pair were co-resolved.

Table 6.15 Pair4: Co-resolution across TRTs

Student Pair		TRT 1	TRT 2	TRT 3	No. / Total LREs (% co-resolved)
		Co-resolved/ Total LREs (%)	Co-resolved/ Total LREs (%)	Co-resolved/ Total LREs (%)	
4	Aki	7/23 (30%)	4/13 (31%)	0/19 (0%)	11/55 (20%)
	Yuta	4*/6 (66%) *1 unresolved	1/10 (10%)	0/8 (0%)	5/24 (17%)

Table 6.16 shows a similar trend across TRTs with regards to extended metatalk: in TRT 1 when revising Aki's transcript 34% (8/23) of the LREs produced had involved extended metatalk, and half (3/6) of those produced during revision of Yuta's transcripts; however, in TRT 3 only 9% (2/21) of the LREs produced during revision of Aki's transcript entailed extended metatalk, and none of the LREs during revision of Yuta's transcript.

Table 6.16 Pair 4: Quality of LRE metatalk across TRTs

Pair		TRT 1		TRT 2		TRT 3		No. / Total LREs (% of Total LREs)	
		E	L	E	L	E	L		
4	Aki	7	15	4	8	2	19	13/55 (24%)	42/55 (76%)
	Yuta	3	3	3	7	0	8	5/24 (21%)	19/24 (79%)
Grand pair total across TRTs								18/79 (23%)	61/79 (77%)
Key: E = Extended metatalk; L = limited metatalk									

Closer examination of the pair-talk data revealed that this pair's trend away from more interactive EWL seemed related how pairs had co-resolved LREs in TRTs 1 and 2. TRT 1 began with revision of Aki's transcript, and the 'co-resolution' of this pair's first LRE is illustrative. In Excerpt 23 below, in Turn 4 Aki suggests changing the lexical item 'one' into 'the first'. Yuta pointed out that the revision is not necessary

(Turn7) and Aki explained that his suggested revision was for the sake of style (Turn 10). Aki's revision also entailed adding the definite article before 'first' and 'that' after. Yuta responded (Turn 11) that "so many" changes were not needed, but Aki insisted on making this change to his speech.

Excerpt 23

- 4 Aki I HAVE THREE REASONS: ONE THE STUDENTS WILL ahh... I
 want to change this to 'Firstly' or 'First'. First, Second, Third.
- 5 Yuta Hmm? First, second, third?
- 6 Aki Yes, sorry, will cut 'one' and use word 'First'. 'The first is'.
- 7 Yuta I think you need not use 'First' or 'Second, Third' in this. Your
 speech is ok, if you say 'One, two'. We can understand this.
- 8 Aki I have three reasons the short summer.
- 9 Yuta One, two, three.
- 10 Aki But I think, *ma*, 'first' 'second' sounds better style. 'The first
 The first is that.
- 11 Yuta We don't need to make so many changes, its same.
- 12 Aki But, I want to...let's change it.

While Yuta was willing to constructively contribute to revision of errors; of the seven 'co-resolved' LREs, four concerned stylistic changes Aki had wished to make to his transcript, *every one* of which was disputed by Yuta on the grounds that they were not needed. These four LREs were coded as 'co-resolved' in that they had entailed deliberation; however, these LREs were ultimately resolved by Aki insisting on the revisions over Yuta's objections. Although difficult to capture on by transcription, on the recordings Yuta began to sound increasingly exasperated with each of Aki's suggested stylistic changes, and Aki mildly exasperated that Yuta kept disputing them. When the pair turned to revising Yuta's transcript, Aki also suggested making two 'style revisions'. Yuta refused to make one of these suggested revisions and only grudgingly made the other.

The same pattern emerged in TRT 2 which also began with revision of Aki's transcripts. Of the four LREs coded 'co-resolved', two concerned Aki again wishing to make changes for the sake of style. Again, Yuta expressed feeling these revisions were unneeded. Later, while revising Yuta's transcript, Aki once more suggested a stylistic change, as shown in Excerpt 24 below. Yuta responded to the suggestion (Turn 64, below) by saying "*mendokusai*", which politely translated means 'what a hassle' or less politely as 'what a pain', and declined to make the revision. Whether Yuta was referring to Aki himself as being 'a pain' or merely his suggestion, the result was Aki ceased to initiate any further revisions to Yuta's transcript in TRT 2.

Excerpt 24

- 61 Aki 'And'? How many times you say 'and'?
62 Yuta In this paragraph? Uh...I should count?
63 Aki You don't use too many 'and'? Maybe cut this and make
'next' or hmm 'also'.
64 Yuta *Mendokusai*, it is not needed to change these, it's not a
problem.
65 Aki <9> Ok.

In the final TRT (TRT 3)—which also commenced with Aki's transcript—the two LREs that involved extended metatalk were both Aki talking aloud to himself. In TRT 3 all but two of the revisions were self-initiated and resolved by the student whose transcript was being revised. Although both students minimally attended to the revisions the other put forth verbally—insofar as the revisions were verbally acknowledged and written into their copy of the other's transcript—by the final TRT session the pair appeared to have reached an unspoken *modus vivendi* of revising their own transcripts largely independently.

6.3 Results for Research Question 1b: EWL during final-revision

Here results are reported regarding the outcomes and nature of the LREs produced by pairs during final-revision of transcripts where pairs compared their revisions to

those the teacher had made on the teacher-revised copy of the same transcripts. The class had been able to identify 43.36% (219/505) of inaccurate language items contained in the transcripts. Learners noticed 48.81% (165/338) of inaccurate grammar forms, 26.51% (35/132) of inaccurate lexical items, and 54.28% (19/35) of the sentences/clauses in need of reformulation for accuracy. Note that recurring tokens of an identical error were counted as a single error. For example, Ken had misused 'poors' (i.e., 'poor people') four times in TRT 3, and these 4 tokens were counted as a single lexical error. If each token of an identical error is counted separately, the class noticing of noticing of total errors drops to 39.24% (219/558). Taken as a class, the learners in this study noticed fewer points in need of correction than earlier TRT studies where pairs had noticed 60% (Lynch, 2001) and 52.35% (Stillwell *et al.*, 2010) of language points in need of correction.

As Table 6.17 shows, within the class there was considerable variation between pairs in the percentages of errors noticed.

Table 6.17 Noticing of inaccurate language across TRTs

Pairs and Students	TRT 1	TRT 2	TRT 3	Total noticed / Total errors (%)
	No. Noticed / Total error (%)	No. Noticed / Total error (%)	No. Noticed / Total error (%)	
1 Naoto Ken	13/33 (39.39%)	19/59 (32.20%)	17/40 (42.50%)	49/132 (37.12%)
2 Chika Momo	21/38 (55.26%)	16/42 (38.09%)	23/39 (58.97%)	60/119 (50.42%)
3 Naho Asami	18/38 (47.36%)	10/45 (22.22%)	21/49 (42.85%)	49/132 (37.12%)
4 Aki Yuta	18/43 (41.86%)	17/32 (53.12%)	26/47 (55.31%)	61/122 (54.46%)
Grand total noticed/Total errors (%)				219/505 (43.36%)

Clear indications of a connection between pair EWL and the percentage of errors noticed were not found in the above data. For example, Naoto and Ken (Pair 1), despite having spent 20 minutes revising every transcript twice, only noticed on

the same percentage of errors (37.12% or 49/132) as Naho and Asami (Pair 2) who had spent only 10 to 12 minutes revising each of their transcripts once. The unnoticed form-errors produced by these two pairs were similar, both had a similar proportion of lexical errors (35.82% for Pair 1 and 34.21% for Pair 2) which all learners noticed relatively less frequently, and the majority of remaining errors were in verb morphology/subject-verb agreement, noun pluralisation, conjunctions, and articles. Similarly, the data could not account for the drop in Pair 3's noticing of from 47.36% in TRT 1 to 22.22% in TRT 2; nor account for the rise again in noticing to 42.85% in TRT 3. The types and proportion of errors were similar across the TRTs and there nothing in Naho and Asami's pair-talk data to indicate a different degree of EWL during TRT 2. Using Pair 4 as another example, Aki and Yuta's predominantly limited metatalk and social disengagement in TRT 3 did not lead to a lowering in the percentage of noticed errors; in fact, it was during TRT 3 that these learners noticed the highest percentage of their errors (albeit of their own individual errors).

The primary focus of this section, however, is how pairs engaged with language when reading the teacher's revisions. Relative to a pair's revisions, there were three types of teacher revisions:

- Type 1: The teacher- and pair-revision of a language feature were identical.
- Type 2: Teacher-revision of a language feature that the pair had not discussed as being in (potential) need of revision for accuracy.
- Type 3: The teacher- and pair-revision of a language feature were different or the pair had made a revision the teacher had not.

No cases were found of Type 1 teacher-revisions generating extended metatalk; rather, students merely acknowledged that their revisions and the teacher revisions had been the same (e.g., "Oh good, same"). This was unsurprising because, once it was confirmed that both revisions were identical, there was nothing more for the students to process. In fact, it was common for Type 1 not to be directly

mentioned at all as students were looking for differences between transcripts noticed.

Uptake of Type 2 teacher-revisions was universal with no examples found of learners not adopting a teacher's revision of a previously unnoticed inaccurate language feature. Again, this was perhaps unsurprising as the students were being shown that their output had been inaccurate and uptake was coded as having a positive (-/+) outcome.

Type 3 revisions, in contrast, required pairs to decide whether to adopt the teacher revision or stay with their initial-revision of the same language feature. Here, uptake/adoption of a teacher-revision could either result in an (-/+) outcome in cases where the pair's initial revision had been inaccurate, or result in an (+/+) outcome in cases where the initial-revision had been equally acceptable. A decision to reject of a teacher-revision could also have an (+/+) outcome in cases where a pair was correct in deciding that their initial-revision was also acceptable, or have a negative (-/-) outcome when such decisions were incorrect. Differences in how pairs discussed Type 2 revisions and how they arrived at their decisions whether or not to adopt Type 3 teacher-revisions are discussed below.

6.3.1 Pair 1: Naoto and Ken and Pair: 2 Chika and Momo

Type 2 revisions: As shown in Table 6.18, for Pair 1, 41% (34/83) of all discussions of Type 2 teacher-revisions of previously unnoticed errors entailed extended metatalk, with higher rates for episodes addressing lexical revisions (52% extended) or teacher reformulations (60% extended), than revisions to form (33% extended). In Pair 2, 56% of all LREs were extended, and over 50% regardless of whether the revision had been lexically-, form-, or reformulation-focused.

Table 6.18 Pairs 1 and 2: EWL with Type 2 teacher-revisions

Pair 1: Naoto and Ken							
Lexical revision (n=29)		Form revision (n=49)		Reformulation (n=5)		Total LREs (n=83)	
Metatalk		Metatalk		Metatalk		Metatalk	
E	L	E	L	E	L	E	L
15/29 (52%) TC = 3	14/29 (48%)	16/49 (33%) TC = 1	33/49 (67%)	3/5 (60%) TC = 1	2/5 (40%)	34/83 (41%) TC=5	59/83 (59%)
Pair 2: Chika and Momo							
Lexical revision (n=22)		Form revision (n=35)		Reformulation (n=2)		Total LREs (n=59)	
Metatalk		Metatalk		Metatalk		Metatalk	
E	L	E	L	E	L	E	L
12/22 (55%) TC = 2	10/22 (45%)	20/35 (57%) TC = 10	15/35 (43%)	1/2 (50%) TC = 1	1/2 (50%)	33/59 (56%) TC=13	26/59 (44%)
Key: E = Extended; L = Limited; TC = Teacher consultation							

Episodes of limited metatalk regarding form-revisions were usually clearly cases of the teacher having revised a type of error the learners had merely inadvertently overlooked, as illustrated in Excerpt 25 from Pair 1 during TRT 3 (teacher-revisions in underlined **bold**):

Excerpt 25

- 79 Naoto (reading teacher-revision) BUT A PRIVATE SCHOOL IS A
PROFIT MAKING ORGANIZATION.
- 80 Ken Two 'a', missed it.

Limited discussions of teacher-revision of lexis also seemed to be cases of what appeared to be for the students, in retrospect, 'obvious' errors, as illustrated below in Excerpt 26 from Pair 1, TRT 1:

Excerpt 26

- 34 Ken (reading teacher-revision) THEY EASILY FORGET WHAT
THEY LEARNED IN SPRING SEMESTER.

35 Naoto (groans) Not 'lose memory'. So simple (laughs).

Original: [...]THEY EASILY LOSE MEMORY WHAT THEY LEARNED [...]

Episodes of extended metatalk were always attempts to determine themselves the reason for a teacher's revision. In cases of such discussions of lexical revisions, additional teacher consultation was rare as the pairs could usually find satisfactory explanation by consulting their dictionaries, as shown in Excerpt 27 (Pair 2, TRT 1), below:

Excerpt 27

27 Chika IF YOUR SUMMER VACATION IS NOT LONG ENOUGH, YOU
CAN'T HAVE SUFFICIENT, what is this?

28 Momo 'Sufficient'?

29 Chika Ah ummm ah, it must be different from 'abundant'. Let's
check.

30 Momo You do 'sufficient' I will 'abundant' <25> (silence while
checking dictionaries) show me 'sufficient'.

31 Chika Here says *jubun na*.

32 Momo This ('abundant') is *hofu na*. Mmm his (the teacher's
revision) is better for your idea.

33 Chika Yeah, because I want to say 'enough' time (Note: 'enough'
and 'sufficient' both translate as '*jubun na*').

Original: [...] YOU CAN'T HAVE ABUNDANT TIME TO DO INTERNSHIP [...]

Between Pairs 1 and 2, the main difference in extended metatalk over teacher-revision of form was that, when discussing form revisions, Naoto and Ken (Pair 1) more often felt able to account for the revisions without teacher consultation. Only 1/12 of Pair 1 discussions of previously unnoticed form-errors involved consulting the teacher; whereas half (10/12) of Pair 2's extended discussions had involved the teacher. In large this was due to Pair 1's previously unnoticed inaccurate forms more often being relatively easily explained errors in use/absence of articles, noun pluralisation, subject-verb agreement, or word form (e.g., 'feels equality' → 'feels equal.')

Chika's transcripts—were often more difficult aspects of English grammar. Excerpt 28 below from TRT 3 provides an example:

Excerpt 28

- 43 Momo **COMPARED** TO PUBLIC SCHOOLS, it's different.
- 44 Chika Comparing to...compared to...comparing to...compared to
(sounding the alternatives aloud)
- 45 Momo (laughing) Me too.
- 46 Chika (laughing) Why? I think its ok 'comparing'.
- 47 Momo Maybe...I don't know but is 'comparing' mmm 'comparing' a
domeishi (gerund)?
- 48 Chika No, I don't think that. It's a verb, right? I mean, I am
comparing public and private schools. I need to ask this
(signals teacher). Question. Why just...did you change to
'compared to'? Why is 'comparing' wrong?
- 49 Teacher Chika where are we here? What line?
- 50 Chika Line 6, I know you want to say 'you compare this school
with that school', but why not 'comparing'?
- 51 Teacher Yeah, I...no, it's because 'Comparing' and 'to' do not go
together. If you want to keep the word 'comparing',
because you love '-ing', you would have to say 'comparing
public schools to private schools'.
- 52 Chika Ahh (confused?).
- 53 Teacher 'Comparing', you really need something, an object. Like this
<15> (writes on board) if you use 'comparing' you need A
and B, like 'Comparing apples and bananas, apples are less
expensive so blah, blah, blah.'
- 54 Momo Too long (laughs).
- 55 Teacher Well, not so long, I just changed to 'compared to' to fix it
- 56 Chika faster.
- 57 Teacher So, to use 'comparing'?
If you love 'comparing', you need to write the longer 'A to B'
pattern like this (indicates example on board). It would be
'Comparing private schools to public schools' uh, where...
- 58 Chika WE GET BETTER EDUCATION.
- 60 Momo Oh, oh, oh (I see). Got it. (teacher leaves)

- 61 Chika How to change it?
'Compared.' I don't want to write more words, so just his way.

Original: Comparing to public schools we get better education by going to private school.

Revision: **Compared** to public schools we get better education by going to private school.

Type 3 revisions: As Tables 6.19 and 6.20 show, both pairs noticed 23 Type 3 revisions where the teacher had revised language differently from the students or had not revised something the pair had. (Again, recurrent correction of an identical error was counted as a single revision noticed repeatedly for determining engagement with language revision.) In both pairs, 61% of their decisions whether or not to amend their initial-revisions had entailed extended metatalk, and the pairs declined to adopt the majority of Type 3 teacher-revisions. These decisions were by-in-large reached without teacher consultation, and only three decisions (all made by Pair 2 in episodes of limited metatalk) had a negative (-/-) outcome (i.e., the learners had failed to recognize that their revision had been inaccurate).

Table 6.19 Pair 1: Metatalk and uptake of Type 3 teacher-revision

Pair 1 (Naoto and Ken): metatalk and uptake				
No. revisions and LREs	Limited metatalk 9/23 (39%)		Extended metatalk 14/23 (61%)	
	Adopted	Declined	Adopted	Declined
23	4	5	6 (TC=1)	8 (TC=2)
Uptake outcomes				
Adopted 10/23 (43%)		Declined 13/23 (57%)		
(-/+) 4	(+/+) 6	(+/+) 13	(-/-) 0	
Key: TC = Teacher consultation				

Table 6.20 Pair 2: Metatalk and uptake of Type 3 teacher-revision

Pair 2 (Chika and Momo): metatalk and uptake				
No. revisions and LREs	Limited metatalk 9/23 (39%)		Extended metatalk 14/23 (61%)	
	Adopted	Declined	Adopted	Declined
23	1	8	4 (TC=0)	10 (TC=6)
Uptake outcomes				
Adopted 5/23 (22%)		Declined 18/23 (78%)		
(-/+)	(+/+)	(+/+)	(/-/)	
1	4	15	3	
Key: TC = Teacher consultation				

In episodes of extended metatalk, the two reasons articulated for declining to adopt a teacher's revision were: (a) they felt the pair- and teacher-revisions were functionally equivalent or the formal difference between revisions was irrelevant, and (b) cases where the learners felt the teacher's revision did not capture the speaker's intended meaning as accurately as their own (studies of uptake of written teacher feedback have also reported feedback being 'rejected' when seen as altering their intended meaning ; e.g., Storch and Wigglesworth, 2010; Swain, 2006).

Figure 6.1 contains an excerpt of Pair 2 discussing the difference between the teacher's and their revision, and coming to the conclusion that the teacher's revision was no better than their own.

Figure 6.1 Example of learners declining to adopt teacher-revision (Equivalency)

Original [LINE:16]	AND IN SUMMER VACATION WE CAN REFRESH AND CHARGE OUR MOTIVATION FOR NEXT SEMESTER.
Pair-Revisions	AND IN SUMMER VACATION WE CAN <u>BE REFRESHED</u> AND CHARGE OUR MOTIVATION FOR NEXT SEMESTER.
Teacher-Revisions	AND IN SUMMER VACATION, WE CAN REFRESH <u>OURSELVES</u> AND CHARGE OUR MOTIVATION FOR NEXT SEMESTER.
Final-Revisions	AND IN SUMMER VACATION, WE CAN <u>BE REFRESHED</u> AND CHARGE OUR MOTIVATION FOR NEXT SEMESTER.
Pair-talk: Final-revision (TRT 1) (M = Momo [author]; C = Chika [revision partner]; T = Teacher)	

101	M: WE CAN REFRESH OURSELVES . Hmm (thinking).
102	C: (laughs) He just always changes different...never worry, no worries. Just keep going.
103	M: But why did he change?
104	C: Uh, the difference? He like just didn't think to use it, uh adjective.
105	M: But 'ourselves'? It is 'us'?
106	C: Yeah, 'refresh us' is same idea, like ah his wasn't adjective, how say <i>tadōshi</i>
106	(<i>transitive verb</i>)?
107	M: I don't know English but I see. So go? (i.e., continue reading)

Figure 6.2 contains an excerpt of Pair 1 declining to adopt the teacher's revision of inserting the conditional marker 'when' in lieu of their inserted marker 'if'. Ken (Turn 25) wondered whether both revisions were conditional markers. Naoto explained that, indeed, these both revisions indicated conditions, but that the teacher-revision expressed a certainty that students forget what they had learned before summer vacation (Turn 26). Ken, jokingly, noted that the teacher's revision may actually be a more apt description (in his case at least) of the negative effects of long summer vacations (Turns 27-31). Naoto, however, declined the revision on the grounds that he wanted to express more of a possibility than a certainty (Turn 32).

Figure 6.2 Example of learners declining to adopt teacher-revision (Difference)

Original [Line: 7]	AND THEIR MEMORY OF SPRING SEMESTER ARE LOST, THEY HAVE A GAP BETWEEN SPRING SEMESTER AND AUTUMN SEMESTER.
Pair-Revision	<u>IF</u> THEIR MEMORY OF SPRING SEMESTER ARE LOST, THEY HAVE A GAP BETWEEN SPRING SEMESTER AND AUTUMN SEMESTER.
Teacher-Revisions	AND, WHEN THEIR MEMORY OF SPRING SEMESTER IS LOST, THEY HAVE A GAP BETWEEN SPRING SEMESTER AND AUTUMN SEMESTER.
Final Revisions	AND, (1) <u>IF</u> THEIR MEMORY OF SPRING SEMESTER (2) IS LOST, THEY HAVE A GAP BETWEEN SPRING SEMESTER AND AUTUMN SEMESTER.
Pair-talk: Final-revision (TRT 1) (N = Naoto [author]; K = Ken [revision partner])	
24	N: AND WHEN , ok, 'WHEN' is different.
25	K: It's not a condition?
26	N: Yes, uh, no. A condition but his means students always lose their memory.
27	K: That's true (laughs) like me.

- 28 **N:** (Laughs)
 29 **K:** I cannot speak English.
 30 **N:** Because... (humorously drawn out)
 31 **K:** I forget last time semester.
 32 **N:** (Laughs) But I think 'if' is closer to my meaning.
 33 **K:** Not always forget memory. Keep?
 34 **N:** Mm (yes). Let's (continues reading)

In episodes of extended metatalk, Pairs 1 and 2 rarely articulated that their reason for *adopting* a teacher revision was that they were worried their own revision might be formally inaccurate *per se*; indeed, in only 8/46 cases where the teacher's and pairs' revisions differed had the pair-revisions been inaccurate. Rather, these learners mainly adopted teacher-revisions because either (a) they felt the teacher-revision captured their intended meaning more clearly, or (b) most often because they were attracted by the novelty of the teacher-revisions. Figure 6.3 contains an excerpt of Pair 1 adopting a teacher-revision for reasons of novelty.

Figure 6.3 Example of learner adoption of teacher-revision (Novelty)

Original [Line:11]	THE PURPOSE OF EDUCATION IS TO RAISE SOCIAL USEFUL PEOPLE.
Pair-Revisions	THE PURPOSE OF EDUCATION IS TO RAISE <u>USEFUL PEOPLE FOR SOCIETY</u> .
Teacher-Revisions	THE PURPOSE OF EDUCATION IS TO RAISE <u>SOCIALLY</u> USEFUL PEOPLE.
Final Revisions	THE PURPOSE OF EDUCATION IS TO RAISE (1) <u>SOCIALLY</u> USEFUL PEOPLE.
Pair-talk: Final-revision (TRT 1) (K = Ken [author]; N = Naoto [revision partner])	
72	N: THE PURPOSE OF EDUCATION IS TO RAISE <u>SOCIALLY USEFUL</u> =
73	K & N (in unison) 'socially useful'!
74	N: Same as us I think.
75	K: But his is shorter. I did not know 'socially'.
76	N: <i>Kakkoi</i> (<i>cool</i>). He made it <i>fukushi</i> (<i>adverb</i>).
77	K: I like this (teacher-revision) more. Change it to it.

In episodes of limited metatalk it was not possible to unequivocally establish the reasons why learners declined or adopted a teacher-revision. However, most of the teacher's 'revisions' declined with limited metatalk had been teacher *non*-revision of language these pairs had amended for stylistic reasons (e.g., to increase the lexical variety of conjunctions). Student directives such as "skip it", "just ignore", or "never mind this" likely indicated that the pairs recognized that the teacher had simply not made those types of stylistic revisions. Pair-adoption of revision with limited metatalk were all cases where during initial revision the learners had already deliberated between two revision options (e.g., 'at midnight' versus 'in midnight') and the teacher's revision merely showed that they had chosen the wrong option.

In sum, the above excerpts illustrate that Pairs 1 and 2 appeared to have been just as highly engaged with the reading of teacher-revisions as they had been with the making of their initial-revisions. The pairs often used reading of the teacher's revisions—especially when revisions differed—to reflect upon their own and each other's use of language and often such reflections included discussion of issues of nuance of meaning. Additionally, as during initial-revision, the features of laughter and humor in the pair-talk seemed to indicate that Pair 1 and 2 had enjoyed the process of working together to make final-revisions as much as they had making initial-revisions.

6.3.2 *Pair 3: Naho and Asami*

Naho and Asami, in stark contrast to Pairs 1 and 2, showed very limited EWL when making their final revisions. Table 6.21 shows that in only 13 out 83 cases (13% of cases) where the teacher had revised previously unnoticed errors was extended metatalk produced by this pair. In the vast majority (87%) of cases, the revisions were merely read, repeated, and written into their copies of a given transcript.

Table 6.21 Pair 3: Metatalk and Type 2 teacher-revisions

Lexical revision (n=22)		Form revision (n=55)		Reformulation (n=6)		Total LREs (n=83)	
Metatalk		Metatalk		Metatalk		Metatalk	
E	L	E	L	E	L	E	L
2/22 (9%) TC = 1	20/22 (91%)	8/55 (15%) TC = 3	47/55 (85%)	1/6 (17%) TC = 1	5/6 (83%)	11/83 (13%) TC = 5	72/83 (87%)
Key: E = Extended; L = Limited; TC = Teacher consultation							

Excerpt 29 below is an example from TRT 2 of one of the rare occasions when this pair discussed a teacher's revision extensively. The teacher's revision concerned the need to invert the verb in sentences that begin with a negative adverb phrase (in this case: 'Only when...can we...'). As can be seen below, this inversion caused the pair considerable confusion. In Turn 66, Asami noted that 'Only when' was a "*kateiho*" (condition) and therefore her original 'we can' was correct (which would be true if the sentence utterance had began with 'When...'). Naho (Turn 69) speculated that perhaps all conditional sentences required "*touchi*" (inversion); a grammar term Asami appeared unfamiliar with, and which Naho (Turn 71) explained. Ultimately, the pair found the revision inexplicable and beckoned over the teacher for explanation.

Excerpt 29

- 62 Asami ONLY WHEN WE ARE COLLEGE STUDENTS CAN WE GO eh?h?
'can we'?
- 63 Nao "can we"? "we can"?
- 64 Asami "we can" is ok maybe.
- 65 Nao Your sentence is correct...correct too, maybe. WE CAN WEAR OUR PRIVATE CLOTHES.
- 66 Asami 'can we' I don't understand. I think 'can we' is the grammar. 'Only when' is *kateiho* (conditional).
- 67 Nao *Kateiho*, ok.
- 68 Asami But then 'can we'. I don't know why rule.
- 69 Nao 'Only when' is *kateihou* (conditional) so next sentence (i.e. clause) need this uh maybe mmm to make *touchi*? (word

- inversion)*
- 70 Asami *Touchi?*
- 71 Nao Mmm, make this (i.e., 'we can') *hantai (reversed)*.
- 72 Asami I don't <5> my grammar is poor, became poor. We need him (i.e., the teacher) (beckons the teacher).
- Original:* Only when we are college students we wear our private clothes.
- Revision:* Only when we are college students can we wear our private clothes.

The learners' sophisticated use of (Japanese) grammar terminology in the above excerpt indicated that it was not a lack of vocabulary needed to discuss teacher revisions that had prevented the pair from more often discussing revisions extensively. Therefore, the possibility needs to be entertained that the reason these learners so infrequently engaged in extended metatalk was that, in the majority of cases, the reasons the teacher had made his revisions had been transparent for these learners.

This pair's limited EWL was also evident in cases where teacher's and the student's revisions differed (Type 3 revisions). As shown in Table 6.22, in the overwhelming majority (79%) of cases the learners did not discuss these revisions extensively, and adopted 79% of the teacher's revisions. Interestingly, in only one case had their own revisions been inaccurate (i.e., only one decision to adopt a teacher-revision had an (-/+) outcome). That is, with one exception, *not* adopting a teacher revision would have had a positive outcome, but the pair almost never discussed the possibility that their own revisions could be accurate (nor ever asked the teacher). Pair 4's adoption of teacher-revision without discussion might have been a sign of Asami and Naho having had more confidence in their native-speaker teacher's English than their own; as Stillwell *et al.* (2010) speculated was the case with their learners. This possibility is strengthened by the fact that two episodes of extended metatalk had been deliberation over whether to *erase* a revision they had made to an item the teacher had not.

Table 6.22 Pair 3: Metatalk and uptake of Type 3 teacher-revisions

Pair metatalk and uptake				
No. revisions and LREs	Limited metatalk 11/14 (79%)		Extended metatalk 3/ 14 (21%)	
	Adopted	Declined	Adopted	Declined
14	11	0	0	3 (TC=0)
Uptake outcomes				
Adopted 11/ 14 (79%)		Declined 3/ 14 (21%)		
(-/+)	(+/+)	(+/+)	(-/-)	
1	10	3	0	
Key: TC = Teacher consultation				

6.3.3 Pair 4: Aki and Yuta

As shown in Table 6.23, noticing of teacher-revision of previously unnoticed errors only produced extended metatalk in 21% (13/61) of episodes among Pair 4. Here too, it was perhaps the case that most teacher-revisions did not generate extended discussion because the underlying reasons revisions were transparent to the learners.

Table 6.23 Pair 4: Metatalk and Type 2 teacher-revision

Lexical revision (n=24)		Form revision (n=34)		Reformulation (n=3)		Total LREs (n=61)	
Metatalk		Metatalk		Metatalk		Metatalk	
E	L	E	L	E	L	E	L
7/24 (29%) TC = 2	17/24 (71%)	4/34 (12%) TC = 2	30/34 (88%)	2/3 (67%) TC = 2	1/3 (33%)	13/61 (21%) TC=6	48/61 (79%)
Key: E = Extended; L = Limited; TC = Teacher consultation							

Table 6.24 shows that, 57% of Aki and Yuta's decisions whether or not to amend their initial-revisions in light of teacher-feedback were reached without extended metatalk, and there were only four cases of the teacher having been consulted. In-line with Pairs 1 and 2—but unlike Pair 3—Aki and Yuta did not adopt

the majority of teacher-revisions: 64% (18/28) of these revisions were declined, and all but one of these decisions had positive outcomes.

Table 6.24 Pair 4: Metatalk and uptake of Type 3 teacher-revision

Pair metatalk and uptake				
No. revisions and LREs	Limited metatalk 16/28 (57%)		Extended metatalk 12/28 (43%)	
	Adopted	Declined	Adopted	Declined
28	6	10	4 (TC=3)	8 (TC=6)
Uptake outcomes				
Adopted 10/28 (36%)			Declined 18/28 (64%)	
(-/+)		(+/+)	(+/+)	(-/-)
4		6	17	1
Key: TC = Teacher consultation				

While the percentage of Pair 4 LREs which had included extended metatalk during initial-revision had declined over the course of study (section 6.2.4); the percentage of episodes of extended metatalk when making final-revisions remained constant over the course of the study. For example, across TRTs 1 to 3, only 20-22% of discussions of previously unnoticed errors included extended metatalk.

However, the nature of the (albeit infrequent) extended episodes did change over the study. In TRT 1, Aki and Yuta extensively discussed the reasons for the teacher's revisions only three times, but one of these was Yuta pointing out that the teacher had not made a 'stylistic' revision that Aki had suggested. This episode is shown in Figure 6.4, below.

Figure 6.4 Pair 4 discussing teacher non-revision of style

Original [Line:2]	I HAVE THREE REASONS: ONE THE STUDENTS WILL NOT STUDY [...]
Pair-Revision	I HAVE THREE REASONS: <u>THE FIRST IS THAT</u> STUDENTS WILL NOT [...]
Teacher-Revision	(None)
Final Revision	I HAVE THREE REASONS: <u>THE FIRST IS THAT</u> STUDENTS WILL NOT [...]

Pair-talk: Final-revision (TRT 1) (A = Aki [author]; Y = Yuta [revision partner])	
6	A: I HAVE THREE REASONS: ONE THE STUDENTS WILL NOT STUDY, TWO]
7	Y: He did not change.
8	A: Hm? (pardon?).
9	Y: This, he didn't make change 'first'. Yours was not wrong. We don't need to change.
8	A: Ah (mild laugh) change is ok, mmm not wrong. My style. It's better.
9	Y: But...he didn't this... we can forget style.
10	A: It is ok to change style. Want to ask?
11	Y: No. Read.

In the above excerpt, Aki and Yuta are rehashing the arguments for and against making 'stylistic' revisions to their transcripts they had made when making their initial-revisions during this TRT (see section 6.2.4). Yuta appeared to take the teacher's non-revision as support for his view that stylistic revisions need not be made. Aki (Turn 8) seemed to take this all in good humor and explained it was just his choice to improve his transcript stylistically. When Yuta pressed his point (Turn 9), Aki even suggested asking the teacher (Turn 10), but Yuta declined (Turn 11). However, during this revision of Aki's transcript, Yuta continued to note cases where the teacher had not made one of Aki's stylistic revisions (e.g., "It's not changed"), and also when the pair revised Yuta's transcript. While the pair did not discuss their differences of opinion further, it is difficult to see how Yuta's continued noting of the teacher's non-revisions could have positively contributed to pair-dynamics.

The point here is that the above, plus Yuta's having described Aki's stylistic revisions as being "a pain" earlier in TRT 2 (see section 6.2.4), appeared to have negatively affected Aki's willingness to further engage with Yuta when reading teacher-revisions. In TRT 2, whenever Yuta posited a reason to explain a revision the teacher had made to his own transcript, Aki responses were non-committal (e.g., "ok"). When Aki's transcript was being revised, episodes of extended metatalk were of Aki beckoning the teacher and asking the teacher to confirm his ideas regarding

whether to amend his initial-revisions; that is, those episodes were discussions with the teacher but not Yuta. This pattern continued in TRT 3 where as previously described each learner had for all intents and purposes revised their own transcript.

Aki and Yuta's interactions clearly illustrated, as Storch (2004) observed, that peer collaboration requires learners to have shared or compatible task goals. Pair 4 clearly disagreed on the parameters of the TRTs: Yuta felt revision should be restricted to correction of errors, whereas Aki took a more expansive view of revision which additionally encompassed making changes to language for the sake of style. As the author Yuta was in his rights to decline making stylistic revisions to transcripts of his own speeches, but he also objected to such revisions being made to Aki's transcripts. Why Yuta felt the two differing views on the parameters of revision to be *incompatible* was unclear, but perhaps his unspoken goal was simply to complete the task as quickly (or to be more generous, 'efficiently') as possible.

6.4 Summary of results

6.4.1 Summary of noticing during initial pair revision of transcripts

Across three TRTs, the eight learners produced a total of 290 LREs when pairs revised transcripts independently from the teacher (initial-revision). The frequency with which learners notice language features can be indicative of learners' states of alertness, and therefore are a possible measure cognitive EWL (Svalberg, 2009). As means of promoting learner noticing of inaccurate language in their oral output, TRTs in this study were found to have been less successful than in earlier studies. Although there was considerable variation between the four pairs; overall, when making initial-revisions the eight learners in this study only noticed 43.36% (219/505) of inaccurate language items contained in the transcripts, compared to previous TRTs where pairs had noticed 60% (Lynch, 2001) and 52.35% (Stillwell *et al.*, 2010) of language points in need of correction. However, the learners in this study noticed lexical items in need of revision more frequently than did learners in

previous TRT studies: 22.14% of revisions were to lexical items compared to only 7% and 1% of revisions made by learners in Lynch (2001) and Stillwell *et al.* (2010) respectively. As in previous TRT studies (Lynch, 2001; Stillwell *et al.*, 2010) learners' revisions were not always of errors: 22.41% (65/290) of revisions concerned changes to already accurate language, often for stylistic rather than corrective reasons.

6.4.2 *Summary of learner EWL during initial pair revision of transcripts*

A perhaps better measure of cognitive EWL than rates of noticing, however, is how learners discussed the language features they noticed in transcripts, something previous TRT studies have largely not addressed. While Lynch (2001) and Mennim (2012) provided select episodes of extended learner discussion of language during TRTs, neither specified how frequent such episodes were. As a gauge of the cognitive effort put into making revisions, on the whole this study found TRTs had not generated a high proportion of episodes that involved extended metatalk. Although there was considerable variation between pairs (summarized section 6.4.3, below), 65.17% (189/290) of learners' discussions of language involved limited metatalk where one learner supplied a revision which was then acknowledged and written without further discussion. Extended metatalk was only involved in 34.82% (101/290) of LREs. The data also revealed that extended pair-talk was not required for learners to resolve most LREs correctly: 95.23% (180/189) of episodes of limited metatalk were resolved accurately compared to 77.22% (78/101) of episodes of extended metatalk. Actual co-resolution of LREs was even less frequent, 'co-resolved' LREs being those where learners had deliberated and resolved a language problem by 'talking it through' (Swain and Lapkin, 2002) to co-construct a revision. In each pair, only 20.25% - 23.28% of LREs were co-resolved.

The quality of pair metatalk was found to be influenced by the type of language errors learners noticed. Revisions of grammar forms were by far the most common type of revision made but the least likely to generated extended

discussion: 68.62% (199/290) of LREs had addressed form, but only 30.15% (60/199) were extended. In contrast, clauses and sentences in need of wholesale reformulation for clarity generated the most extended metatalk—69.23% (18/26) of reformulations involved extended metatalk and most frequently entailed co-resolution—but these revisions were the least common and the focus of only 8.96% (26/290) of LREs. Revision of lexis also generally generated a relatively high proportion of extended metatalk, especially discussions over word choice and meaning—46.15% (30/65) of discussions of lexis were extended—but only 22.41% (65/290) of LREs were lexically focused.

In sum, results did not show initial pair revision of transcripts to be especially effective in promoting extended engagement with language or co-resolution, nor found such extended engagement to be necessary for learners to correctly revise most of language features noticed. Although there were significant differences in EWL between pairs (summarized 6.4.3, below), the universal ability among learners to accurately repair such a large proportion of their inaccurate language without extended metatalk strongly indicates the likelihood that a significant amount of the inaccurate language learners produced, especially grammatical inaccuracies, had not been what Corder (1974) labeled ‘errors’ but rather ‘mistakes’ or ‘slips’. An *error* is an inaccurate use of a form arising from a gap in the learner’s understanding of the TL system; whereas, a *mistake* is an inaccurate use of a L2 form that a learner has explicit knowledge of but has not yet fully mastered, and *slip* is an accidental slip of the tongue (Corder, 1974 cited in Ellis and Barkhuizen, 2005, p. 64). When the learners had delivered their relatively impromptu speeches, the need for real-time delivery/communication meant they had to rely more on their intuitive ‘implicit’ L2 knowledge which is only available through automatic processing (Ellis, 2009, p. 12). During TRTs, learners had the time to explicitly focus on form and could more easily draw upon their declarative ‘explicit’ L2 knowledge (Ellis, 2008) which has generally been found to be more accurate than a learner’s implicit knowledge (e.g., Loewen,

2009). Given these learners' ability to accurately revise most of the inaccurate grammar forms they noticed without extended metatalk makes it likely that many had been performance-pressure related mistakes (in Corder's sense) rather than errors. This would explain why extended discussion language, especially of discrete grammar items, was found to be so relatively infrequent: when self-noticed or brought to the learner's attention by their partner, mistakes were easily self-corrected; when corrected by a peer, the nature of the mistake was obvious and the revision accepted without need for further discussion.

6.4.3 Summary of differences in EWL between pairs

Table 6.25 summarizes the main similarities and differences in EWL among the four pairs in this study during initial revision of transcripts.

Table 6.25 Comparison of pair EWL during initial revision of transcripts

Pair No.	% Extended LREs	Signs of positive (+) /negative (-) social or affective Engagement	Time taken to finish initial revision
1	41.09% (30/73)	(+) Laughter; jokes; compliments; thanks for assistance	40 min.
2	45.67% (37/81)	(+) Laughter; jokes; compliments; thanks for assistance; affective support (e.g., 'you can do it').	40 min.
3	28.07% (16/57)	(?) Absence of overt signs of enjoyment of task but also no explicit mentions of dislike.	20-24 min.
4	22.78% (18/79)	(-) Disagreement over task goals (-) Unwillingness to take direction (-) Unwillingness to interact	TRT 1 = 40 min. TRT 2 = 33 min. TRT 3 = 32 min.

I will begin with a recap Pair 4's (Aki and Yuta's) interaction, as they were the only pair whose interaction changed over the study. As described in sections 6.2.4 and 6.3.3, an inability to agree on task goals and the merits of stylistic revisions led to a move from interactive and otherwise collaborative interaction (TRT 1) to non-collaboration and minimal interaction, and increasingly less time spent

completing the task. This was the only pair where the social and affective factors that facilitated or, in this case, impeded EWL were manifest.

Pairs 1 (Naoto and Ken) and 2 (Chika and Momo) evidenced the highest engagement with the task of working together to revise transcripts: these pairs produced the most extended LREs (41%-45.67% of LREs), showed similar signs of positive social or affective engagement between partners, and spent the longest time (40 minutes) to complete initial revision (both pairs, in effect, reading a given transcript twice). In contrast, Pair 3 (Naho and Asami) showed no overt signs of particularly enjoying the task of working together revising transcripts, produced fewer episodes of extended metatalk, and of all pairs, spent the least amount of time on task (20-24 minutes on average). However, it was unclear from the data analysed for this chapter what it was about the task Pairs 1 and 2 found engaging. Similarly, while it was speculated earlier in this Chapter Pair 3 was simply not challenged by task (possibly due to their inaccurate uses of English being easily corrected mistakes), further possible explanations for the differences in EWL between Pairs 1 and 2 versus Pair 3 will be sought from the interview data in Chapter 8.

6.4.4 Summary of Engagement with teacher-revisions

The quality of EWL generated by being provided with teacher-revisions to compare to their own revisions broadly mirrored the engagement a given pair exhibited during the preceding initial revision stage of TRTs. For Pairs 1 and 2, discussions of revisions made by the teacher to language features they had not previously noticed were extended 41% (34/83 episodes) and 56% (33/59 episodes) of the time respectively, a similar proportion of episodes involving extended metatalk as these two pairs had produced during initial revision of transcripts. Noticing of cases where their previous revisions and the teacher's differed generated even greater engagement: both pairs noticed 23 such teacher-revisions in the course of the study,

63% (14/23) of which, in both pairs, generated extended discussions over the differences and whether or not to adopt the teacher's alternative revision.

For Pair 3, noticing of teacher-revisions seldom generated extended metatalk. In only 13% (11/83) of episodes where pairs noticed the teacher had revised language they had not previously discussed involved extended discussion. Such teacher-revisions were merely noted and copied. Cases where the teacher's and Pair 4's revisions differed, the learners overwhelmingly adopted with teacher's revision without discussion, despite their own revisions in the vast majority of cases having also been perfectly acceptable. As for Pair 4, their 'discussions' of teacher revisions were extended in only 28% (25/89) episodes, and as related previously, subsequent to TRT 1, these extended discussions were held mainly between the individual whose transcript was being revised and the teacher; rather than between the learners.

One finding of note concerning learners' engagement with teacher-revisions was the frequency with which learners opted *not* to adopt a teacher's revision when it differed from their own. Stillwell *et al.* (2010) had speculated that the reason their learners had retained and employed more of their teacher's revisions than their own when re-performing oral tasks was that the learners had more faith in the teacher's revisions being correct. Their speculation is plausible. Studies of L2 learners' use of teacher written corrective feedback in redrafts of their compositions have found learners reporting they do have more faith in their teacher's than peer's feedback (e.g., Paulus, 1999; Yang, Badger and Yu, 2006), and also passively accepting and using teacher without understanding its significance (e.g., Hyland, 1998; Lee, 2007; Zhao, 2010). This may have been the case with Pair 3 in this study (see section 6.3.2) who adopted 79% of teacher-revisions in lieu of their correct pair-made revisions without deliberation, but it was not the case with the other learners who discussed

the differences in revisions and ultimately only chose to adopt 36%-43% of the teacher's revisions over their own.

7 CHAPTER 7: ENGAGEMENT WITH LANGUAGE AND RETENTION OF REVISIONS

Research Question 2 asked whether differences in EWL during transcript revision affected retention of revisions. Following Storch (2008) and using the same operationalization of limited/extended metatalk, this chapter reports the results of investigation of the effects two aspects of EWL had on retention of revisions. The two aspects of EWL investigated were: (i) the quality of metatalk (i.e., limited vs. extended metatalk) as a gauge of cognitive engagement, and (ii) the learner's participation in revisions which had been the product of limited metatalk. Whereas Storch (2008) reported on the effect engagement had on retention of revisions made to learners' written text reconstruction task; this chapter reports on the effect engagement had on retention of revisions made to learners' oral output.

7.1 Data sorting for analysis

7.1.1 Sorting of data for analysis of effect of quality of metatalk on retention of revisions

Although learners revised 24 transcripts of initial deliveries, because one student was absent for one redelivery of a speech, revisions made to 23 transcripts of initial speech performances were compared to the language produced in the 23 corresponding redeliveries. The coded data from the 'Retention Results' tables (see section 5.2.2) found 396 final revisions made to transcripts of initial-speech deliveries where there had also been relevant opportunities for those revisions to be used in the corresponding redeliveries of speeches (see section 5.2.2 for examples). Of these 396 revisions, 44 were tokens of revision of identical errors (see section 5.3.6) which were *excluded* from the retention data because it was impossible to isolate the effect quality of metatalk on retention from the effect of multiple noticing of the same error.

Of the remaining 352 final revisions, 278 had been identified as *single round* revisions. That is, either (a) pair-revisions made during the initial rounds of transcript revision which pairs subsequently found to be identical to the teacher's revisions (and therefore left unchanged and not discussed further); or (b) teacher-revisions adopted by pairs during the final rounds of transcript revision that had been made to language the pairs had not previously discussed (i.e., errors pairs had not noticed themselves during initial revision). Every single round revision could be matched to a single LRE coded as either having manifested 'limited' (L) or 'extended' (E) metatalk. If a revision re-appeared in a redelivery of a student's speech, the revision was counted as having been 'retained'; and if it did not appear in the redelivery as 'not retained' (see section 5.5 for examples).

The revision data also included 74 *double round* revisions which concerned language features discussed during both the initial- and final-revision rounds (see section 5.5). These were incidences of pairs noticing that their own revisions and the teacher's revisions of the same language feature had differed and deciding whether or not to adopt the teacher's revision in lieu of their own. Each double round revision had been coded as if the product of two LREs—one per revision round—both concerning the same language point with the LREs coded as either manifesting limited (L) or extended (E) metatalk (again, see section 5.5). However, for the purposes of investigating the effect quality of metatalk had on retention, double round revisions were conceptualized as the product of a single discontinuous LRE (see section 5.2). That is, double round revisions were considered to be qualitatively akin to the discontinuous LREs identified in pair-interaction during initial revision where learners had made a revision to a language feature, but later returned to the feature and re-amended it. Therefore, double round revisions where the corresponding LREs had been coded (E + E), (E + L), or (L + E) were grouped together as revisions having involved extended metatalk. An argument could be made that double round revisions with LREs coded (L + L) should *also* be classified as 'extended'

because choosing between two potential revisions was a criterion for coding metatalk as extended. However, double round revisions coded (L + L), were the product of two LREs lacking this study's other criteria for identifying extended metatalk: the presence of any justification for, or deliberation over, revision options being articulated. All revisions with LREs coded (L + L) were cases of learners passively adopting the teacher's revision without discussion or rejecting a teacher's revision out of hand. Therefore, double round revisions with LREs coded (L + L) and single round revisions coded as only having involved limited metatalk (L) were considered qualitatively similar enough to group together. As with single round revisions, the criteria for counting double round revisions as 'retained' or 'not retained' was whether the learner's final revision re-appeared in redelivered speeches. (As noted earlier in section 5.5, no incidences of initial pair-revisions rather than adopted teacher-revisions, nor cases of rejected teacher-revisions, were found to have appeared the redelivered speeches).

7.1.2 Sorting data for analysis of effect of participation in episodes of limited metatalk on retention of revisions

Following Storch (2008), in this study the single round revisions were examined to investigate whether mere participation in episodes of limited metatalk affected retention of their corresponding revisions. Of the 278 single round revisions identified as having had occasions for use in the corresponding redeliveries of speeches, 204 had been coded as having been the products of episodes of limited metatalk. 'Participation' in episodes (LREs) of limited metatalk referred to whether or not the learner who had produced the language item being revised had either initiated the corresponding LRE and/or resolved it by supplying the revision (see section 5.3.4). That is, to for a learner to be considered to have participated in a single round revision made to his or her transcript, the corresponding LRE had to have been coded:

- *Self-initiated/Self-resolved* (SI/SR)
- *Self-initiated/Other-resolved* (SI/OR)
- *Other-initiated/Self-resolved* (OI/SR)

The learner was considered not to have participated in a revision made to their transcript if the corresponding episode of limited metatalk had been coded:

- *Other-initiated/Other-resolved* (OI/OR)
- *Teacher-initiated/Teacher-resolved* (TI/TR)

7.2 Results

7.2.1 *Effect of quality of metatalk on retention of revisions*

A total of 350 revisions (278 single round + 74 double round) made to transcripts of 23 speeches were identified as having had occasions for use in the redeliveries of speeches. Every speech had been redelivered one week after it had been revised. As Table 7.1 below shows, overall, learners retained 65% (228/352) of their transcript revisions when they redelivered their speeches. The table also shows that 74% (98/132) of revisions that had involved extended metatalk were retained versus only 41% (90/220) of revisions that involved limited metatalk. These results were in-line with the descriptive statistics reported in Storch (2008) and Storch and Wigglesworth (2010) who also found a greater percentage of language amendments (made to written texts) that were the product of 'elaborate/extensive' LREs being retained one week later versus amendments which had been the product of episodes of limited metatalk.

Going beyond descriptive statistics, a chi-square test of independence with a Yates correction factor was performed to examine whether the relation between quality of metatalk and retention of revision was significant (alpha: $p < .05$). The

relation between these variables was found significant. The chi-square statistic with Yates correction (1, $N = 352$) was 7.64. The p -value was .0056 and significant at $p < 0.05$. This finding further bolsters Storch (2008) and Storch and Wigglesworth (2010) contention that revisions which are the product of limited metatalk are less likely to be retained than revisions which entail explicit justification/explanation or deliberation.

However, while the effect of quality of metatalk on retention of revisions was found to be statistically significant, the overall effect of quality of metatalk on retention was quite modest. Table 7.1 reveals a difference of only 12.5 revisions between the number of revisions that learners actually retained (the **Observed** frequency) and the number of revisions learners would have been retained if retention were purely a matter of random chance (the **Expected** frequency). In other words, out of a total of 352 revisions made, only 12-13 *more* were retained due to having been the product of extended metatalk than would have been retained by mere chance. (Conversely, only 12-13 fewer revisions were retained due to having been the product of only limited metatalk than would have been by chance). Again, the effect of quality of metatalk on retention, while statistically significant, was not dramatic.

Table 7.1 Quality of metatalk and retention of single round revisions

	Retained		Not Retained		<i>Row Totals</i>
	Observed	(Expected)	Observed	(Expected)	
Extended metatalk	98	(85.5)	34	(46.5)	132
Limited metatalk	130	(142.5)	90	(77.5)	220
Column Totals	228		124		352 (Grand Total)

7.2.2 *Participation in limited metatalk and retention of revisions*

Out of 278 single round revisions, 204 had been the product of limited metatalk. As Table 7.2 below shows, 51 of these revisions had involved the learner whose transcript was being revised participating in the revision (either by noticing the language feature and/or supplying the revision. The remaining 153 revisions had been entirely the work of either the revision partner or the teacher and involved no participation on the part of the learner whose transcript was being revised beyond merely copying the revision into his/her transcript.

Merely participating in episodes of limited metatalk appeared to effect retention of revisions. Table 7.2 shows that 80% (41/51) of such revisions in which the learner had participated were retained by the learner, whereas only 52% (80/153) of revisions in which the learner had not participated were retained. These descriptive statistics were in-line with those reported in Storch (2008); however, note that in her study 'participation' entailed the learner to have both initiated the noticing of language *and* provided the revision.

Another chi-square test of independence with a Yates correction factor was performed to examine whether the relation between participation in episodes of limited metatalk and retention of revision was significant (alpha: $p < .05$). The relation between these variables was found significant. The chi-square statistic with Yates correction was 11.38. The p -value was .0007 and significant at $p < 0.05$. Therefore, merely repeating and writing a revision that had been initiated and supplied by another (i.e., having no participation in the making of the revision) was found to lead to retention less often than participating in the making of revisions involving limited metatalk.

However, as Table 7.2 shows, while the effect of participation in episodes (LREs) of limited metatalk on retention was found to be statistically significant, the overall effect of participation versus non-participation was also found to be modest. Comparing observed frequency of retention of revisions to that expected by mere chance, the difference between the observed and expected frequencies was only 10.75 revisions. That is, out of 204 revisions, only 10-11 *more* revisions were retained due a learner having participated in corresponding episodes of limited metatalk than would have been expected by chance. (Conversely, only 10-11 fewer revisions were retained due to non-participation than would have been by chance). In sum, the effect of participation in episodes (LREs) of limited metatalk on retention, while statistically significant, was also found not to be dramatic.

Table 7.2 Participation in episodes of limited metatalk and retention of revisions

	Retained		Not Retained		<i>Row Totals</i>
	Observed	(Expected)	Observed	(Expected)	
Participation	41	(30.25)	10	(20.75)	51
No participation	80	(90.75)	73	(62.25)	153
Column Totals	121		83		204 (Grand Total)

7.3 Summary of results

Of the 352 revisions made to 23 transcripts of initial speech transcripts where relevant opportunities to have used those revisions one week later in the 23 corresponding redelivered speeches, 65% (230/352) had been retained (including both single round and double round revisions). The study's criterion for retention was that a revision reappeared exactly as made (with the exception of the presence of dysfluencies which were disregarded). The only prior TRT study with which to compare is Lynch (2007) who reported that his 12 learners had produced 53%

(25/47) 'right' uses of revisions in repeat oral task performances 1 week after performing TRTs.

The results from the study's statistical analysis found that extended metatalk (i.e., which included deliberation or explanation of language) had resulted in statistically significant, superior short-term learning gains (i.e., retention of revision) than limited metatalk in which deliberation or explanation was absent. The analysis further found that participation in episodes of limited metatalk resulted in statistically significant, superior retention of revisions than merely repeating and writing a revision that had been initiated and supplied by another. However, the statistical data also indicated that the *overall* effect of both quality of metatalk and participation in episodes of limited metatalk had on retention of revisions was very modest. When observed and expected frequencies of retention were compared, for quality of metatalk, the difference between observed and expected frequencies was a modest 12-13 (12.5) revisions; and for participation in episodes of limited metatalk, the difference between frequencies was a mere 10-11 (10.75) revisions.

Note, however, these findings do *not* imply that inaccurate language produced during initial speech performance would have been spontaneously reproduced accurately in the second speech performances if the oral tasks had merely been repeated without the intervening TRTs. All studies investigating the effects of task repetition I am aware of (Ahmadian, 2011; Bygate 1996, 2001; Bygate and Samuda, 2005; Gass *et al.* 1999) have failed to find any statistically significant effect for mere task-repetition on linguistic accuracy of repeat task performances. The purpose of this study was not to investigate whether revision *per se* had an effect on L2 learning. Rather, this study investigated whether the extent learners discussed language features being revised, and whether or not learners noticed and/or revised language themselves, had an effect on retention of revisions made. The short answer to both questions was 'Yes, but not very much'. The implications

of this finding for use of TRTs as language learning activities will be discussed in Chapter 9.

8 CHAPTER 8: Learner Perspectives on Transcript Revision Tasks

Research question 3 asked how learners perceived Transcript-Revision-Tasks as language learning opportunity and experience. In the post-course interview, all eight students reported that this had been their first and only experience of listening to, transcribing, and correcting their oral English performances. As this had also been my first time as a teacher using TRTs, the primary purpose when designing the interview instrument was to see to what extent the students had enjoyed the tasks and found them beneficial as L2 learners. This chapter examines how students said they felt about (a) hearing and self-transcribing their oral performances; (b) working with a peer while editing transcripts; (c) being asked to speak English only when editing with partner; (d) re-delivering their speeches; and (e) students' recommendations for improving implementation of Transcript-Revision-Task Cycle in future iterations of the course. The interview data collected also revealed insights into their beliefs about language learning, learning from peers, and how these beliefs had served as facilitators or impediments to their engaging with language during TRTs.

8.1 Student perspectives on self-transcription

The first section of the interview instrument (Items 1 and 2: see *Appendix E*) elicited students' views on the benefits and difficulties of self-transcription. This section will first address the difficulties and negative aspects of self-transcription followed by the perceived benefits.

8.1.1 Negative aspects of self-transcription

Student responses largely assuaged concerns that they would find being asked to self-transcribe their recordings overly onerous. All eight participants said that producing verbatim transcripts, ranging from 367 – 629 words in length, required 40-60 minutes to complete, and none of respondents felt this to be an undue amount of homework, or as Ken put it, "I have more longer homework another

classes, so that's not bad." However, two less positive aspects of transcribing were raised. While the majority (seven of eight participants) did not find transcribing particularly difficult—for example, Chika commenting, "Not a little difficult, never was, so yeah, it wasn't, ah, difficult at all," or Naho noting, "My record is, I play it very slowly, so not difficult with *softo* (i.e., transcription software)"—Aki expressed having had more difficulty stemming from pronunciation problems:

Something was difficult, uh...my pronounce is very bad and sometimes I can't listen the...listen what I say. So I play much again. (Aki)

In addition, participants raised one other 'negative' aspects of self-transcription. Three students (Aki, Yuta, Chika) reported *feeling embarrassed* when they listened to recordings their 'poor' English performance, as exemplified by Chika's comment below:

It's also like ah...a little, I mean a little...stressful for me to do that homework sometimes yeah?, because I know last time all I said early, like, are horrible. But I have to finish it [i.e., transcribing] and listen to it. But I know actually it helped me...but...yeah... I was....I feel like ah...embarrassed. (Chika)

8.1.2 Benefits of self-transcription

Despite the negative aspects raised, the eight students were nonetheless unanimously supportive of the idea of being recorded and self-transcribing in future English courses, as exemplified by the comments below:

Teacher should use this strategy, I think. So I'm happy to be recorded in future class. In this class, I did feel nervous, but

as I said it gave me objective point of view and that was very useful. (*Naoto*)

Yes, I want more record. That will be useful (*Yuta*).

What emerged clearly was that students perceived the self-transcription tasks as being *beneficial* ('useful', 'helpful') for learning. These perceived benefits are summarized in Table 8. 1. Comments in support of self-transcription fell into four main themes: (a) raised awareness of actual oral ability; (b) listening alone being insufficient for noticing errors; (c) the opportunity to focus on form not available while delivering speeches; and (d) goal setting. It is important to note that these four themes overlapped each other. For example, the first theme, 'raised awareness of actual oral ability' may have been said to have stemmed from finding 'errors only noticed when transcribing,' which may also have then lead to 'goal setting'.

As previously mentioned, this was the first time for participants to have ever listened to themselves speaking English, as, for example, related by Momo:

Ah....ummm...when I was high school or junior high school students, I don't have opportunity of recording speaking English. So, this university's English class is first time. So I didn't know my English speaking skill (*Momo*).

As Table 8.1 below indicates, one of reasons for finding self-transcription 'useful', raised by all eight students, was that students had been unaware of their actual speaking ability, which all participants related as having noticed as having been *worse* than they had thought (e.g., Comments 1, 2, and 3: Table 8.1, below). Most students also expressed having been 'surprised' by this realization.

Table 8.1 Positive aspects of self-transcribing own oral performance

Positive aspects [No. of respondents]*	Exemplifying comments
Raised awareness of actual oral ability [8]	1. I feel something surprised. I recognized that the English skill is bad more than I think before (<i>Aki</i>). 2. Useful, I think. I learned hard before, before high school so was surprised how down is my English skill so I know...ah...I know my level (<i>Asami</i>). 3. Shock! Like....I would be like, "No, I didn't want to say it". I don't think I talk like this, but I said it. (<i>Chika</i>).
Transcription increases noticing of error [8]	4. And...only voice or recording is ah, not enough. I think eh....if I write what I say, if I don't have the paper, I can't see mistakes so I don't study enough. So transcribe is...has meaning. More powerful (<i>Naho</i>). 5. I think by writing my... what I said I can know my skill clearly...more...clearly. So writing what I said is important, I think. Just..oh....listening is umm shortage? '<i>tarinai</i>' [insufficient] (<i>Asami</i>).
Provides opportunity to focus on form not available during oral performance [6]	6. Umm...in our class, first speech class, maybe I couldn't listen to my English or think grammar, so it is good we record our voice and in my home we can listen that. In speaking, I... my thinking is just saying idea, I can't....read my speaking. When I listening recording...I had...I think I should use that word or not. I do like this (slaps forehead) (<i>Ken</i>). 7. Ah.....when I was speaking in English, my brain is...etto... panic. And speak is difficult for me. Ah...when I was speaking I focused on the next word, not before....before speech, what I said before, so....I didn't notice I was....I make a mistake? But recording is memory...a record, so I can hear my voice and speaking with calm?...uh, without panic (<i>Momo</i>).

Allows for goal setting [3]

8. I hear that, ummm, because my vocabulary was very small: 'and, and, and'; 'but, but'; 'so, so'; so I wanted to use more variety words (Aki).

9. While you are, like transcribing your first speech, then you can, like... sometimes it will help you...it will help you....like ah...what kind of English point you need to improve and what kind of like ah....um....like ah....while you do the next debate you can try to improve those things. Transcribing can, uh, make me try improve my points (Chika).

** Note: Since some students raised multiple themes in their responses, the number of occurrences does not correspond exactly to the number of participants.*

In addition, all eight students expressed the idea that, self-transcribing was 'more powerful' than only listening (Naho: Comment 4, above) and that just listening to the recordings was 'insufficient' for noticing errors (Asami: Comment 5).

Six students further commented that they had found it *difficult or impossible to focus on form during oral performance*, and therefore appreciated the opportunity that listening to the recordings provided. For example Ken reported (Comment 6) that when delivering speeches he couldn't 'listen to' his English or 'think grammar' and rather that "my thinking is just saying idea, I can't....read my speaking." Momo (Comment 7) commented that the pressure ('panic') of real-time delivering of speeches made it impossible for her to monitor her English, unlike when listening to recordings of her oral performance.. of that when speaking, she had to focus on saying 'the next word' and not what she had 'said before' and therefore was unable to notice her mistakes. This theme was perhaps most clearly articulated by Naoto who related that when speaking:

That time, I have no room for checking how often the bad habits appear or how smoothly I speak or best

vocabulary... like that. I can just notice those things when I listen the recording. (Naoto)

Three students (Aki, Chika, Ken) raised the additional point that an increased awareness of their oral performance lead to *goal setting*. For example, Aki (Comment 8) explained wanting to increase the range of vocabulary he uses, specifically citing the limited range of conjunctions he typically employs. As Chika (Comment 9) put it, “Transcribing can, uh, make me try improve my points.”

8.2 Perceptions of peer-editing of transcripts

While students were unanimously supportive of self-transcribing their oral performance, the perceived benefits mostly related to *self-noticing*. In contrast, student opinions diverged greatly with regards to the merits of peer-editing of transcripts. Interview Item 3 elicited student views on the advantages and disadvantages of editing transcripts with a peer. Additionally, Item 6 presented students with a proposed (two-week) TRT-Cycle alternative to the three-week Cycle employed in the study. This alternative plan would require students to self-edit and read teacher’s corrective feedback without a partner (see Appendix G). Students were asked whether they would recommend the alternative plan for next year. This alternative, labelled the *NO PARTNER PLAN* (ostensibly designed to elicit ‘advice for the teacher’) also required participants to reiterate or expand upon their views of peer editing and allowed the researcher to check for consistency of participant’s opinion.

8.2.1 Student preferences for future editing with or without a partner

The interview responses demonstrated that among the eight participants, four said would prefer to continue to edit with a partner while three stated they would prefer self-editing, and one student was undecided. In sum, student preferences for future peer- versus self-editing divided almost equally. These pairings and the individual student’s editing preferences are shown in Table 8.2.

Table 8.2 Editing pairs and individual student editing preference

Pairings	Student's respective editing preference
1. Chika and Momo	Peer-editing/Peer-editing
2. Natoto and Ken	Peer-editing/Peer-editing
3. Naho and Asami	Self-editing/Self-editing
4. Aki and Yuta	Undecided/ Self-editing

As can be seen above, these student preferences divided between the editing pairs who, as illustrated in Chapter 6, had evidenced higher degrees of EWL (Pairs 1 and 2) and those who had evidenced lesser degrees of EWL (Pairs 3 and 4). It is worth noting at the outset that because participants had worked with the same (albeit self-selected) editing partner for every TRT, their perceptions on the merits of and preference for peer-editing of transcripts were strongly colored by their having done so with a particular partner. This is particularly important to keep in mind regarding Aki and Yuta's comments as this pair had evidenced significant discord when revising transcripts.

8.2.2 *Reasons for rejecting the NO PARTNER PLAN and retaining peer-editing*

The four students in favour of peer-editing had only positive things to say about the experience and firmly rejected the *NO PARTNER PLAN*. The perceived benefits of, and reasons for, retaining peer-editing are summarized in Table 8.3, below which fell into four overlapping themes.

The first theme was simply that students felt having a partner led to the noticing of more errors. This reason was given by all four students in favour of retaining peer editing. For example, Naoto (Comment 10: Table 8.3) felt a partner

helped him to notice habitual errors that tended to slip by him; and Momo (Comment 11) felt her partner was helpful in identifying Momo's 'complex' errors.

The second theme was that editing with a partner provided them with opportunities for new learning. Specifically, that having a partner exposed students to a variety of ideas, language forms, and expressions that might not have occurred to them had they been editing alone. This idea was expressed by all four students in favor of continued peer-editing when specifically asked about the value of having a partner to help edit "your transcript". For example, Chika (Comment 12) stated that she had made her partner's corrections her own. Ken (Comment 13) felt that self-editing only led to one possible correction whereas a partner can provide alternative corrections and that together they can produce a more 'creative' final revision. His comment could aptly serve as an alternate to Swain's (2006) oft cited definition of collaborative dialogue. These four students also believed it was useful "for you" to help edit a partner's transcript. Despite having delivered a speech taking the same position supported by the same brainstormed reasons, the students felt helping to correct the partner's speech also exposed them to 'different' or 'new' ways 'of expression' (see: Comment 14).

Table 8.3 Reasons for rejecting *NO PARTNER PLAN* and retaining peer-editing

Reasons [No. of respondents]*	Exemplifying comments
More errors noticed [4]	10. Because partner has a objective point of view, I think. There is... uh... individual people have unique tendency of making mistake, ummm... and those kind of mistakes tend to be slipped by oneself, by myself (<i>Naoto</i>). 11. I read my transcription. And I didn't notice the hard mmm complex? mistakes. But if there is Chika correct...ah....she notice this mistake and correct, yes (Momo).

Opportunities for new learning [4]

12. Yeah, maybe she [i.e., her partner: Momo] didn't realize this but...sometimes I made her corrections and some ideas to make my own. I think it's good point to have a partner (*Chika*).

13. I say that I and Naoto talking and discuss and understand each other's speech. And we use....ah...we make one answer, like many idea but together make final answer. That is better answer, I think. So...this [NO PARTNER] plan don't have a creativity enough, I think (*Ken*).

14. Yes I think so. Ahm...because the speech I made and the speech Ken made has same content but the expression are different. And...Yusuke's version gives me a chance to read another expression which means the same idea as I said, but the different words. (*Naoto*).

Enhances teacher feedback [4]

15. The teacher's edit is ah...one type of answer but is only one. This not limited because we can discuss and discuss and we think hard if can make a... another way to answer or just me copy teacher (*Ken*).

16. My partner helped me understand how corrections are different, I mean sometimes your correction is very different and he helped me understand it (*Naoto*).

17. I think useful. Because I tend to think teacher revision is best, so I almost my transcription just copy teacher revision. But with my partner....my partner say 'your expression is right too', so partner helped my choice. (*Momo*).

More social [3]

18. If you have time, it is enjoy to have a partner, more social (*Momo*).

19. Like, ah, this [NO PARTNER] plan is too much, alone time? So it's kind of fun to have a partner to check it (*Naoto*).

** Note: Since some students raised multiple themes in their responses, the number of occurrences does not correspond exactly to the number of participants.*

The third and closely related theme (also expressed by all four students) was that having a partner when reading the teacher's corrective feedback enhanced the teacher's feedback. One reason was that having a partner exposed students to more possible editing options than they would encounter if restricted to self-editing plus teacher feedback (see: Comment 15, for example). Naoto added the notion that having a partner is helpful sometimes to better understand the teacher's feedback (Comment 16). Interestingly, Momo felt that without her partner she would not have had confidence in her equally acceptable revisions and have adopted the teacher's (Comment 17). She went on to add that had been a good thing as she felt her own revisions would be easier to remember than the teacher's: "Maybe, in the future, expression from my brain is easy to use...easier to remember, more easy". (Momo)

The fourth (final) reason for rejecting the *NO PARTNER PLAN*, expressed by three of the four students, was simply that the 'social' and 'fun' aspects of peer-editing would be lost (Comments 18 and 19). In sum, four of eight students rejected the *NO PARTNER PLAN* and unwilling to sacrifice the perceived benefits of peer-editing: namely, the ability of two to notice more errors, the increased opportunities for learning, the enhancement of teacher feedback, and peer-editing's social aspect. As Chika put it:

It [*NO PARTNER PLAN*] is missing something, I think. But I think it [peer-editing] more like ah....yeah, it's much more effective than this [*NO PARTNER PLAN*].

8.2.3 *Reasons for recommending self-editing of transcripts*

In contrast, three of eight students (Asami, Naho, Yuta) recommended the NO PARTNER PLAN with one (Aki) was undecided, but none had found peer-editing to be beneficial. The reasons given, however, were difficult to summarize in table form because those from Naho and Asami (the Pair 3 TRT partners) were different from the points made by Aki and Yuta (Pair 4). Therefore their reasons are summarized by pair directly in the text below.

8.2.3.1 *Naho and Asami*

These learners had three reasons for supporting the *NO PARTNER PLAN* (i.e., *not* recommending peer-revision of transcripts). These reasons were: (i) students will only notice the same errors; (ii) students can always ask the teacher for help; and (iii) there is no need to discuss teacher feedback.

(i) Can only notice same errors: Asami and Naho felt that, while they helped each other notice errors; they were only able to notice the same type of errors:

Naho corrections and mine are same I think. If, eh, I notice her mistake it is not different because, um same mistake? I learn no...not new things. (*Asami*)

We learned same English in high school...so we know how to tran...correct the same. (*Naho*)

Asami expressed the idea that because her partner's mistakes were the same as her own, she learned nothing new from helping Naho. In the interview she further elaborated that while Naho had helped her notice mistakes, these were just inadvertently 'missed things'. In Naho's comment above, we can see that she felt that she and her partner noticed and corrected the same errors in the same way because of their shared educational background. Interestingly, when asked, Naho also went on to explain that this would be the case even with a different Japanese

partner because everyone in the class had studied for the same university entrance ('center') exam:

We have almost same idea. We all study English in high school for center [exam]. So we have studied...maybe we have studied almost same. So we correct...almost same idea. (*Naho*)

(ii) Can always ask the teacher: Unlike their 'pro' peer-editing counterparts, neither Naho nor Asami raised the idea that having a partner enhanced teacher feedback. In fact, both felt it unnecessary to have a partner to look at the teacher's corrections with on the grounds that if there was something they could not understand, "we can always ask you [the teacher]" (*Naho*).

(iii) No need to discuss teacher feedback: Asami and Naho further felt that teacher feedback left them little to talk about. As it came from a native speaker, the teacher's feedback was perceived to be *ipso facto* more 'right' or 'natural' which left little to talk about before adopting the teacher feedback. As articulated by Asami:

Without Naho? uh...I think it is OK to look at teacher's correction only myself. When...when...looking at teacher's ummm...I can...ah, I didn't have anything what we should speak about. You are native speaker so ...so when looking at your feedback we don't have, um anything to talk. Are you wrong? we don't think this. Your correction is enough because it is right...will be right. I think so...eh...um...so I think we don't have to point at your corrections together. (*Asami*)

Naho and Asami's comments indicated that their shared perceptions of the limited benefits of pair-revision had served as an impediment to EWL during

transcript revision. There is some merit to their point that learners notice the same types of errors. If, as speculated in earlier in section 6.4.2, many of the errors students produced were merely ‘slips and mistakes’, then these mistakes would in many cases have been similar; such as missing articles, lack of agreement between subject-verb, or missing genitive ‘s’. Furthermore, such mistakes would have been easily recognized by one or the other student and ‘corrected in the same way’. However, Naho’s assertion that, due to a common educational background, different Japanese students of comparable proficiency will not notice different types of errors is patently wrong. There were a number of episodes contained in the TRT performance data from Naoto and Ken (e.g., Excerpt 16, p. 106) or Chika and Momo (e.g., Excerpt 21, p. 112) of one student noticing a type of error in their own or partner’s speeches that the other had missed. But to be fair, Naho’s assertion may be more of a speculation stemming from her only having had experienced TRTs when partnered with Asami. That these two learners had both felt that there was ‘nothing to discuss’ when looking at teacher-revision because these ‘must be right’, revealed that this pair did indeed have more faith in their teacher’s revisions than their own which explained why these revisions had been adopted without deliberation during their TRTs.

8.2.3.2 Aki and Yuta

These learners had only one reason in common for recommending the *NO PARTNER PLAN*, that students can notice most errors themselves. Yuta, furthermore, believed supplied a second reason that students’ errors are not comparable. A third reason, unique to Aki, was (iii) that peer-revision is only effective if your partner is engaged in the process.

(i) Students can notice most errors on their own. The only reason for recommending self-editing Aki and Yuta had in common was that they felt they could have *noticed most errors without a partner*, as, for example in Yuta’s comment below:

Uh, I think, um...checking the transcribe in the next week is no...not important. It isn't most important. Making the transcribe is more important. Listening the speech, listening the speech is, it can make you know your, your speaking skill and correct and uh...where you're wrong the speech and the pronounce. I can do. Don't need partner these things. (*Yuta*)

(ii) Students' errors are not comparable. Another reason for recommending self-editing from Yuta was that because students make different errors, they cannot use the corrections made to a partner's transcripts for their own benefit. In other words, he did not see a personal benefit in helping to edit their partner's transcripts: For example, Yuta expressed that because he and his partner's 'way of speaking' were not comparable, helping to edit the partner's transcript was, "more for partner than for me":

You can't compare what you said and partner said uh... people have different way of speaking. So another student...ummm...I, I...ummm, is not my mistake. I try to correct the people speak but it...it is useful, uh, it's more for partner than for me. I don't use his English, so it is not good for helping me. (*Yuta*)

(iii) Partner needs 'passion'. Aki, ultimately expressed being undecided regarding the benefits of peer-editing. He articulated the caveat that peer-editing is only useful if one's partner is engaged by the process. While Aki acknowledged that peer-editing could be useful in such a case, he felt that his partner had not been engaged in the revision of Aki's transcripts:

It's useful to have a partner if the partner thought about your speech how to make it better or what problem my speech. If he thought carefully. But uh, more....was low impact for me to have partner. My partner was low impact. (Aki)

For the same reason, Aki also had found trying to peer-edit Yuta's transcript frustrating due to Yuta's 'lack of passion' regarding the editing of his own transcripts:

My partner did not have passion for English, I think. He did not think his...deeply about his speech. I suggest to correct, but he was not very...much thinking. Just 'ok' he says. (Aki)

Aki's perception that his partner had 'lacked passion' about correcting Aki's speech was, in fact, congruent with Yuta's comment that he had felt revising Aki's transcripts had not been "good for helping me". However, to be fair to Yuta, some of what Aki perceived as a 'lack of passion' would have been colored by the fact that he and Yuta had viewed peer-editing as somewhat different activities. In Chapter 6 we saw that Yuta had appeared to view making stylistic changes to transcripts a distraction from what he had seen as the primary task goal rendering transcripts *linguistically accurate or comprehensible* and had declined to adopt Aki's proposed stylistic revision into his speech. Yuta raised this specific issue when asked what aspects of peer-editing he had not liked:

He [Aki] would try to change my way of expression, sometimes change did not need but Aki would say...uh, 'not elegance' or something. But I want keep my way. Why change? It is enough. The English uh not so simple the...what I said was clearly I think. (Yuta)

While Aki's suggestions were meant to help Yuta improve his speeches, from Yuta's comment above it appears that Yuta may have seen these suggestions as critical of how expressed himself in English. Therefore, some of what Aki saw as Yuta's "not thinking deeply about his speech" was rather perhaps Yuta's desire to retain 'authorial control' and his own individual way of articulating his arguments.

8.3 Using the target-language when editing transcripts

Students had been instructed to use English, rather than Japanese, when editing transcripts with a partner. In part, this request had been made because, when seeking permission from my department head to conduct the research, she had felt that having students spend class time editing transcripts in Japanese was not in-line with the purpose of the course. However, L1 use provides learners with "additional cognitive support" that allows them to "analyse language and work at a higher level than would be possible if they were restricted to sole use of their L2" (Storch and Wigglesworth, 2003 p. 760). In other words, having the TRTs conducted in English may have impeded the students' ability to engage with language. Interview Item 4 elicited students' opinion on being asked to perform editing in the target-language.

All study participants were able to and did conduct TRTs predominantly using 'English only'; however, all participants unsurprisingly expressed having had some difficulty doing so and all admitted to having occasionally resorted to using Japanese. Indeed, for all students, many of their otherwise English LREs were peppered with grammar terminology articulated in Japanese, and student use of Japanese-English dictionaries was ubiquitous. However, when asked if in the future the teacher should continue to explicitly instruct students to peer-edit in English, all eight students advised I do so despite acknowledging that editing in Japanese would be easier. Two reasons were given. Firstly, all eight saw a likelihood that not requesting editing be done in English would lead students to conduct editing primarily in Japanese. For example, Asami who stated:

I think it is good to use a little Japanese only little. But you say like that, then talking will become almost Japanese.
(Asami)

Or Chika, who felt:

Like ah...especially in Japan, don't do that! It's going to destroy us. We will not use English....like ah...no, no more. Don't use it, say that in English class. The whole time lesson would be in Japanese. (Chika)

When students were asked *why* using Japanese the entire lesson would be a bad thing, they all replied that doing so would waste their limited opportunity to communicate in English. As Naho who explained:

In Japanese...in Japan we have, uh, we tend to use Japanese...so the English class is very...very precious to speak English. Because in our life, we don't use English.
(Naho)

Nonetheless, five participants went on to note that some student use of Japanese would be necessary or inevitable:

Ah, but some students don't have enough ability to express, like nuance, how they think in English. I think you should say "Use English." But in case it is difficult, too difficult, I'm just thinking students will use it [Japanese] anyway. (Naoto)

And yet, when asked whether the teacher should be 'relaxed' about the students using Japanese, Naoto reiterated the idea that:

But teachers should...I think teachers should make stress on using English. If you just say, "Try in English but you can switch to Japanese" students easily switch to Japanese.
(Naoto)

In short, students felt that not explicitly requesting peer-editing be done in English or explicitly condoning even limited use of Japanese was a 'slippery slope' which would likely lead students to squander their limited opportunities to speak English by performing the TRT entirely in Japanese.

8.4 Importance of redelivering speeches

Redelivery of speeches had been included in the study's Cycle, in part, for the purpose of ascertaining the factors which appeared to influence retention of revisions; but pedagogically speaking this step could be seen as optional. Interview Item 6 proposed a shorter (only two lesson) potential Debate-Cycle for use in future iterations of the course. This plan would retain peer-editing but dispense with students redelivering their speeches. This plan was roundly rejected by all eight participants, regardless of their views of peer-editing, because they felt that the benefits of redelivery were too great to sacrifice. These benefits fell into four overlapping themes, which are summarized in Table 8.4, below.

Firstly, all students saw redelivery of speeches (and the concomitant debate) as the Cycle's main event (see Comment 20, below) and that redelivery of speeches was their opportunity to practice and employ the revisions made to transcripts of the initial delivery of the speech (see Comment 21).

The second, closely related, theme was that, because students had also transcribed their redeliveries, the students could track their learning progress. This perceived benefit was expressed by five of the eight students, as exemplified in Comment 22.

Table 8.4. Perceived benefits of redelivering speeches

Reasons [No. of respondents]*	Exemplifying comments
Opportunity to practice revisions [8]	20. I think the 3rd lesson is most important in, in all lesson? I think...eh...1st lesson and the 2nd lesson is to get ready for 3rd lesson. It is like a practice to grow up my English. To say speech more clearly (<i>Naho</i>). 21. Ah, second speech is....ah, it is useful because we can use the improved ext....corrected sentence. You can use it in second speech and practice (<i>Aki</i>).
Tracking of learning progress [5]	22. But if you didn't do it [redeliver the speech] actually then, you can't understand which part you actually get improved (<i>Chika</i>).
Smoother debates [5]	23. We can speak more clearly what I say so other person can understand my idea reasons better. Debate is less confusion (<i>Yuta</i>).
Incentive to edit more carefully [2]	24. If you...if you don't use the edited transcript next class, I think that students won't think hard about correction because their idea can't use at second time. So second delivery of speech have big advantage (<i>Ken</i>).

** Note: Since some students raised multiple themes in their responses, the number of occurrences does not correspond exactly to the number of participants.*

The third benefit was that the redelivered speeches were easier for the debate opponent to understand and therefore made the subsequent debating go more smoothly, as stated in example Comment 23. The final benefit, from Chika and also Ken (Comment 24), was that knowing they will have a second opportunity to deliver a speech leads students to edit their transcripts more carefully.

8.5 Final student recommendations

At the end of the interviews, students were asked to choose from one of three Debate-Cycle plans: (1) the *NO REDELIVERY PLAN*; (2) the *NO PARTNER PLAN*; or (3)

the Debate-Cycle employed in this study. After which, the students were then asked if they had any additional recommendations for the teacher.

As we have just seen, none of the students were in favor of a Cycle that dispensed with the redelivering of speeches. Three of the eight students (Asami, Naho, and Yuta) recommended I use the NO PARTNER PLAN and had no additional recommendations to make. The remaining five students recommended I re-employ the study's Debate Cycle. While Momo had no additional advice beyond this, the other four all suggested allowing them to independently revise their speeches while transcribing (which they would send to the teacher) and subjecting this independently revised transcript to the second lesson for peer-review. Aki was the first of these students to recommend the above modifications (although he had earlier in the interview recommended the NO PARTNER PLAN):

I have more smart way. We correct at home, homework and show to partner. Because a many mistakes "um" "ah" "a", "the", "students", I can see myself and partner is help only for making idea expression clear. Teacher read my corrections and makes corrections. Shorter. (Aki)

When I pointed out that his plan was still three-weeks long (as compared to the two-week NO PARTNER PLAN), he replied:

[Laughs] But is more short, time for partner is more short, shorter. We can read teacher's correcting more, more time for advice by you [teacher]. (Aki)

Note that the other three of these four students (Chika, Ken, and Naoto) suggested the same modifications independently of Aki (and each other) and without these modifications being suggested by the interviewer. Chika, Ken, and Naoto also gave similar justifications for the modifications, for example, acknowledging that many of

their errors could be self-noticed and thus save time during peer-editing. Chika further stated as having had to resist the temptation to make such revisions during self-transcription in this study, which confirmed that she had indeed been ‘primed’ to revise her mistakes prior to TRTs (see section 6.2.2).

9 Chapter 9 Discussion and Conclusion

In this study, four pairs of learners were introduced to Transcript-Revision-Tasks (TRTs). The study had three main objectives. The first was to gauge how effective the TRTs had been at fostering learner Engagement with Language (EWL). The second was to investigate whether differences in EWL generated during transcript revision impacted learners' ability to employ their revisions when given the opportunity to repeat their speeches. The final objective was to elicit learners' perceptions of TRTs as learning opportunities in the hope of better understand the factors which had facilitated or impeded learners' EWL. To these ends, the following research questions were posed:

1. What was the quality of learner EWL during transcript revision?
2. Did differences in quality of EWL during transcript revision affect retention of revisions?
3. How did learners perceive Transcript-Revision-Tasks as a language learning opportunity and experience?

For purposes of synthesizing findings, this chapter begins by summarizing the findings for Research Questions 1 and 3 in tandem (i.e., the quality of EWL generated during revision and the factors that appeared to facilitate or impede EWL) followed by a summary of findings for Research Question 2 (whether quality of EWL affected retention of revisions). Next, the limitations of the study are discussed and potential avenues of future research into TRTs suggested. The chapter concludes with the pedagogical implications of the study's findings with regards to the use of TRTs as language learning activities.

9.1 EWL during transcript revision: Impediments and facilitators

This study used Svalberg's (2009, 2012) conceptual framework of 'Engagement with language' (EWL) to construct a picture of what took place as learners were building their language awareness during transcript revision. EWL refers to a learner's combined cognitive, social, and affective states and his/her learning processes in situations where the learner is the agent and language is the object of study (Svalberg, 2009, p.147).

9.1.1 EWL during initial-revision of transcripts

Svalberg (2009, 2012) posited that a learner's state of heightened cognitive engagement manifests in a process of focused reflection and problem solving. The measure of cognitive EWL used in this study was *quality of metatalk* which referred to the extent to which a pair discussed a language point they were revising. Following Storch (2008) and Storch and Wigglesworth (2010), LREs were coded according to whether they had evidenced *limited* (L) or *extended* (E) metatalk. 'Limited' metatalk referred LREs in which a revision was proposed without explanation and immediately accepted/adopted without further discussion. LREs coded as 'limited' were considered to be indicators of lower cognitive EWL. 'Extended' metatalk referred to LREs in which a justification/explanation for a proposed revision was provided and/or in which alternative revisions were proposed and deliberated. Such LREs were considered indicators of higher cognitive engagement as they included evidence of reflection and problem solving. (See section 5.3.3 for examples of 'limited' versus 'extended' metatalk).

Across the three TRTs performed, the four pairs produced 290 LREs during the initial revision of transcripts (i.e., when revising without recourse to teacher feedback). The study's findings regarding learner cognitive engagement and revision outcomes during initial revision of transcripts are summarized in Table 9.1, below.

Table 9.1 Cognitive EWL during initial revision of transcripts

Measure of EWL	Operationalization
Limited metatalk	An LRE in which a suggested revision was accepted without explanation/justification being provided or sought.
Extended metatalk	An LRE in which an explanation/justification revision was provided or sought, and/or alternative revisions were considered and deliberated.
Findings	
1. The vast majority of LREs resulted in accurate revision regardless of quality of metatalk: – 89% (258/290) of all LREs resulted in accurate revision. 2. The majority of revisions were products of limited metatalk: – 65% (189/290) of LREs involved limited metatalk. – 35% (101/290) of LREs involved extended metatalk. 3. Limited metatalk resulted in a higher percentage of accurate revision than extended metatalk: – 95% (180/189) of LREs involving limited metatalk resulted in accurate revisions. – 77% (78/101) of LREs involving extended metatalk resulted in accurate revisions.	

Impediments to EWL: Task factors

One factor which appeared to explain the generally low proportion of extended metatalk seemed inherent to the tasks themselves. By far the most common language features learners noticed were discrete grammar forms: 69% of LREs had addressed grammar but only 30% of these concerning involved extended metatalk. In the main, the grammar features were simple, usually morphological (and only involving one word), and often concerned noun plurals, subject-verb agreement, or changes to verb tense or aspect. As discussed in Chapter 6, these types of grammatical inaccuracies were likely performance-pressure related slips or mistakes rather than errors (Corder, 1974) caused by having had to deliver relatively impromptu speeches. This likelihood was further bolstered by the fact that most of the subject-verb combinations learners produced *were* in agreement and most verbs

in the correct tense. The ease with which learners were able to accurately correct these items also indicates they were not systemic errors.

Additionally, during the interviews, all eight learners expressed some variation of the idea that much of the inaccurate language in their own transcripts noticed by the pair could have been noticed and revised themselves without their partner's assistance (see Chapter 8). Indeed, this was found to have actually been the case when the LREs produced during initial-revision were examined for *patterns of initiation and resolution*. LREs were coded as to whether they had been *self-initiated & resolved*; *other-initiated & resolved*; or had evidenced *joint participation* in a revision. The operationalization of these measures and the aggregate findings regarding the initiation and resolution of the LREs produced by all four pairs over the course of the study are presented in Table 9.2, below.

Table 9.2 Patterns of participation and during initial pair-revision

Measure of participation	Operationalization
Self-initiation & resolution	An LRE both initiated and resolved by the <i>learner</i> whose transcript was being revised.
Other-initiation & resolution	An LRE both initiated and resolved by the <i>partner</i> of the learner whose transcript was being revised.
Joint participation	An LRE that involved <i>co-resolution</i> where pair-members made suggestions, counter-suggestions, and deliberated among options. <i>OR</i> An LRE initiated by one pair member and resolved by the other.
Findings	
1. 48% (138/290) of LREs were self-initiated & resolved. 2. 22% (63/290) of LREs were other-initiated & resolved. 3. 31% (89/290) of LREs involved joint participation: – 22% (63/290) of LREs involved co-resolution. – 9% (26/290) of LREs were initiated by one pair member and resolved by the other.	

As above table shows, in total nearly half (48%) of all revisions had been made by the learners whose transcripts were being revised without *any* involvement on the part of the revision-partner. In addition, only approximately a third (31%) of LREs had involved of joint-participation and less than a quarter (22%) of had been co-resolved: where learners had made suggestions, counter-suggestions, pooled linguistic resources, and deliberated over the best revision to make.

In sum, one key impediment to higher levels of EWL during initial-revision of transcripts was likely that many of the inaccurate language features that learners noticed on transcripts of their oral performances differed from errors they would notice in written compositions. When making written compositions, learners have more time to draw more upon their declarative L2 knowledge (Ellis, 2009), but as discussed in section 6.4.2, oral performances require learners to draw mostly upon their implicit knowledge which is more likely to lead learners to produce slips or mistakes. If learners are required to produce verbatim transcripts of oral performance “warts and all” (Mennim, 2003), these ‘slips/mistakes’, once noticed, may simply not necessitate extended metatalk to revise accurately. When self-noticed or brought to the learner’s attention by their partner, these mistakes are easily self-corrected; when corrected by a peer, the nature of the mistake is obvious and the revision accepted without need for further discussion.

Impediments to EWL: Social and affective factors

As discussed in Chapter 6, the types and proportion of types of inaccurate language items were similar across the transcripts of all eight learners. Therefore, these task factors alone could not adequately account for the notable differences in the proportion of extended metatalk produced and patterns of initiation/resolution found between pairs. These differences are summarized in Table 9.3, below:

Table 9.3 Comparison of pair EWL during initial-revision of transcripts

Measure of EWL	Pair 1 (Naoto/Ken)	Pair 2 (Chika/ Momo)	Pair 3 (Naho/Asami)	Pair 4 (Aki/Yuta)
% LREs with extended metatalk	41% (30/73)	46% (37/81)	28% (16/57)	23% (18/79)
% LREs self-initiated & resolved	46% (34/73)	35% (28/81)	53% (30/57)	72% (57/79)
% LREs other-initiated & resolved	21% (15/73)	35% (28/81)	21% (12/57)	8% (6/79)
% LREs involving joint participation	33% (24/73)	35% (28/81)	26% (15/57)	20% (16/79)
% LREs co-resolved	23% (17/73)	22% (18/81)	21% (12/57)	20% (16/79)
	Greater overall EWL		Less overall EWL	

As the table shows, Pairs 1 and 2 produced a higher percentage of LREs involving extended metatalk than Pairs 3 and Pair 4, which evidenced greater overall cognitive EWL during initial-revision of transcripts. Pairs 1 and 2 also evidenced more frequent participation the revision of their partner’s transcript than had Pairs 3 and 4 (i.e., fewer LREs were entirely self-initiated & resolved). As Svalberg (2009, 2012) posits that, with regards to social EWL, a learner engaged with language is initiating and interactive, Pairs 1 and 2 were considered to have shown greater social EWL overall during transcript revision. Only the percentages of LREs which were co-resolved were nearly the same across pairs.

Analysis of learners’ interview responses indicated that it had been differences in learners’ affective EWL that best explained the differences in cognitive and social EWL during TRTs. Svalberg posits that a learner who is affectively engaged will possess a positive orientation towards the language, the interlocutor, and what they represent. Simply put, the learners in Pairs 1 and 2 all expressed the belief that

that having a partner when revising transcripts increased and enhanced opportunities for learning. The perceived learning benefits articulated by these learners were that peers could identify mistakes they could not and that peers could teach or expose them to new vocabulary and grammar.

In contrast, Naho and Asami (Pair 3) believed that there was ‘nothing new’ they could learn from reading and correcting each other’s transcripts because, as Japanese learners, they shared the same educational background and therefore made the same type of errors. For the same reason, both learners also expressed the idea that they could only notice the same types of errors. Aki and Yuta (Pair 4) both shared the belief that they felt they could notice most of their mistakes without a partner. However, Yuta additionally felt that because students make *different* errors, students cannot make use of the corrections made to a partner’s transcripts for their own benefit. But the biggest impediment to EWL for this pair, as explained in section 8.2.3.2, was their inability to see agree upon whether making revisions purely for the sake of style fell within the confines of the task. Interview data showed this to be a rather complex disagreement involving issues of autonomy and control which ultimately led to these learners to disengage socially and no longer seek or offer each other assistance in the revision process. In sum, the interview data revealed that most the significant factor underlying differences in cognitive and social EWL were learners’ perceptions regarding the merits and parameters of peer-editing.

9.1.2 EWL during final-revision of transcripts

Of the 290 revisions pairs had made to transcripts during initial transcript revision, 22% (64/290) had been made to language that was already accurate; however, the vast majority of such revisions (57 out of 64) were the product of extended metatalk in which learners articulated that the rationale for the revision was stylistic, such as to increase lexical variety, rather than to repair a perceived error. The study found that on average pairs had only been able to identify 43% (219/505) of the inaccurate

language items contained in transcripts. During final-revision of transcripts, when pairs compared their own revisions to those made by the teacher, learners noticed three types of teacher-revisions relative to their own:

- Type 1: The teacher-revision of a language feature and the pair-revision were identical ($n = 151$).
- Type 2: The teacher-revision had been made to a language feature that the pair had not discussed revising ($n = 286$).
- Type 3: The teacher-revision of a language feature differed from pair-revision ($n = 88$).

To measure learners' cognitive EWL during final-revision, the LREs produced were analysed in terms of whether they had evidenced extended or only limited metatalk.

No cases of Type 1 teacher-revisions generating extended metatalk were found: students merely noted their revisions had been the same and many of the Type 1 revisions were never directly mentioned at all. This was unsurprising given that once noticed that teacher- and pair-revisions were identical, there was nothing more for the learners to process/discuss, and also because they were on the lookout for differences in revision. Therefore, only the cognitive EWL evidenced in the 374 pair-discussion of Type 2 and Type 3 teacher-revisions are summarized in Table 9.4, below:

Table 9.4 Comparison of pair EWL during initial-revision of transcripts

Measure of EWL	Operationalization
Limited metatalk	An LRE in which a teacher-revision was immediately adopted and an explanation/justification for the revision was neither provided nor sought.
Extended metatalk	An LRE in which an explanation/justification for a teacher-revision was provided or sought, and/or pairs deliberated whether or not to adopt a teacher-revision.

Findings				
Measure of EWL	Pair 1	Pair 2	Pair 3	Pair 4
% LREs with extended metatalk	45% (48/106)	57% (47/82)	14% (14/97)	28% (25/89)
% LREs extended metatalk <u>Type 2</u> teacher-revisions	41% (34/83)	56% (33/59)	13% (11/83)	21% (13/61)
% LREs extended metatalk <u>Type 3</u> teacher-revisions	61% (14/23)	61% (14/23)	21% (3/14)	43% (12/28)
Key: <u>Type 2</u> = Teacher-revision of language feature the pair had not previously discussed as in need of revision. <u>Type 3</u> = Teacher- and pair-revision of language feature differed.				

As they had during initial-revision of transcripts, Pairs 1 and 2 evidenced notably higher overall cognitive EWL than had Pairs 3 and 4 when examining both Type 2 and Type 3 teacher-revisions.

Once again, the interview data indicated that it was differences in beliefs over whether having a partner enhanced opportunities for learning that accounted for these differences in pair-EWL during final-revision of transcripts. All four learners from Pairs 1 and 2 expressed the belief that talking with a peer about the teacher's feedback enhanced their understanding of that feedback and, in the case of Type 3 revisions, allowed them to better decide whether or not teacher-revisions should be adopted in lieu of their own. In total, Pairs 1 and 2 declined to adopt 67% (31/46) Type 3 teacher-revisions—either by discussion amongst themselves or in consultation with the teacher—with 90% (28/31) of such decisions resulting in a positive outcome (i.e., linguistically accurate revision). In contrast, Naho and Asami's (Pair 3) expressed belief that here was 'nothing new' they could learn about the L2 from each other was also articulated as the reason these learners had additionally felt there had been nothing each other could add when reading the teacher's revisions. These learners

also expressed a shared lack of trust in their own L2 efficacy. These two beliefs likely led to the largely passive uptake (i.e., without discussion or teacher consultation) of teacher-revisions, both in cases of revisions directed at language items they had not previously discussed, but also when their (overwhelmingly accurate) revisions had differed in form from the teacher's. Aki and Yuta's (Pair 4) increasing social disengagement over the course of the study led to the great majority of their episodes of extended metatalk being one-on-one consultations with the teacher with no attempt made to involve their pair-partner.

9.2 EWL and retention of revisions

The findings from the revision-retention data showed that the ability to retain a revision one week later was affected by the quality of metatalk that had been involved in the making the revision. Of the 350 transcript revisions identified as having had occasions for use in the redeliveries of speeches, 74% (98/132) of revisions that had involved extended metatalk were retained versus only 41% (90/220) of revisions that involved limited metatalk. A chi-square test of independence with a Yates correction factor was performed and the effect of quality of metatalk on retention of revisions was found significant: the p -value was .0056 and significant at $p < 0.05$). The results indicated that EWL that led to deliberation or explanation of language when making a revision ('extended metatalk') had been superior in terms of short-term retention-*cum*-learning gains than EWL resulting in 'limited' metatalk that had not generated such discussion. However, while the effect of quality of metatalk on retention of revisions was statistically significant, the overall effect of quality of metatalk on retention was found to be quite modest. A comparison between the observed and expected frequencies of retention found only 12.5 more revisions were retained due to having been the product of extended metatalk than would have been retained if retention were purely a matter of random chance.

Findings also indicated that mere participation in revision episodes of limited metatalk had an effect on retention. 80% (41/51) of single round revisions involving limited metatalk made to a learner's transcript in which that learner had either been the person to have noticed the language in need of revision or the person to provide the revision (or both) were retained by the learner. In contrast, merely repeating and writing a revision that had been initiated and supplied by another was found to less often lead to retention. Only 52% (80/153) of such revisions were retained. Another chi-square test of independence with a Yates correction factor was performed and the effect of participation on retention of revisions was found significant: the p -value was .0007 and significant at $p < 0.05$). Again, however, while the effect of participation on retention of revisions which had involved limited metatalk was statistically significant, the overall effect of participation versus non-participation was found to be quite modest. A comparison between the observed and expected frequencies of retention found only 10.75 more revisions were retained due to learner participation than would have been retained if retention were a matter of random chance.

In sum, both the effects on retention of differences in quality of metatalk and that of participation versus non-participation, while statistically significant, were found in this study to be less dramatic than the descriptive statistics, and those reported in the previous studies of Storch (2008) and Storch and Wigglesworth (2010), might have suggested.

9.3 Limitations of study and avenues of future research

The present study had several limitations worth addressing in future research. As a small-scale study involving only eight students, results cannot be generalized to a larger population or even predictive of future learners taking the same course at the same institution. Although the study sheds light on what it is that learners noticed in

their speaking performances in this particular context, what learners notice in a different context cannot be determined from this study. Indeed, the rates of noticing and LRE production in general were lower than reported in earlier TRT studies. That being said, all classroom based studies of TRTs are likely to be small scale as these tasks are most practical with smaller-sized classes. Future studies of classroom use of TRTs are warranted, however. Earlier TRT studies (Lynch, 2001; Mennim, 2012) provided select episodes of extended learner discussion of language during TRTs without establishing how frequent such episodes were. This study found such episodes to have been more the exception than the rule with my learners. More TRT studies are needed to determine if, for example, I am correct in my suspicion that verbatim transcripts of oral performances are likely to contain a high proportion of slips to errors and thus produce relatively low proportions of extended metatalk. Also, this study was the first to investigate how learners actually discuss teacher feedback provided during TRTs. Whether teacher-feedback directed at language features learners have already noticed and discussed better fall within the learners' ZDP merits further investigation. From a sociocultural theory perspective, such feedback would be more likely to be internalized and therefore effective than feedback directed at errors in learners had not noticed.

It is worth reiterating, as has been noted by other researchers (e.g., Sachs and Polio, 2007; Storch, 2008; Storch and Wigglesworth 2010), coding for engagement is highly inferential and learner verbalization may not necessarily reflect the depth of cognitive processing. Although learners commented on their cognitive engagement during the end-of-semester interviews, videotaping TRT performances would have allowed for use of more immediate stimulated recall protocols (e.g., Mackey, Gass, and McDonough, 2000) to prompt more focused learner-comment regarding their thoughts and feelings when participating in specific learning events. Unfortunately, limited resources (video recording equipment) and lack of student availability for stimulated recall sessions precluded such inquiry.

While the study design allowed learners opportunities to repeat their speeches to investigate whether short-term retention differences could be related to differences in cognitive EWL, no tracking was done to look for evidence of more durable long-term gain. Given the relatively small number of corrections (both pair and teacher) made to any single given students' transcript, and the very different topics of the debates, identifying language items which use of would have been obligatory across speeches would likely have been very small in number (Loschky and Bley-Vroman, 1993). I could have additionally looked for effects of EWL on short-term retention and implicit L2 knowledge gains by having examined the learners' debate performances. The debates would have represented even more spontaneous and less rehearsed oral output. The debates were recorded and may contain language items—especially lexis—that had been the focus of revision in TRTs of the position speeches. Only constraints of time prevented me from pursuing that vein of inquiry, and the extant data collected could be used for such future research.

The use of posttests would have strengthened the study's investigation of learning outcomes. Posttesting would have allowed investigation of retention of revisions for which no opportunities for re-employment had arisen during repeat speech deliveries and also for tracking longer term gains. In principle, such posttests should be oral and designed to measure learners' implicit knowledge/learning of the revised forms: such as elicited oral imitation tests (Erlam, 2009). However, such tests would need to be tailor-made (Swain, 1998) to capture what different learners choose to focus on during transcript revision. As Swain and Lapkin (1998) acknowledge, such tests are difficult to design because of the considerable time pressure they place on researchers to develop relevant test items within the time span of a study. Given that transcribing pair-interactions, identifying LREs, and matching them to revisions, even for a single TRT, took weeks to finish; constraints of time precluded use of tailor-made posttests as a viable research option.

A final limitation was that I did not control for the effect of quality of metatalk on learning gains separately from that of mere task repetition. Although all studies investigating task repetition I am aware of (Ahmadian, 2011; Bygate 1996, 2001; Bygate and Samuda, 2005; Gass *et al.* 1999) have all failed to find statistically significant effects for task-repetition on accuracy, a possible future avenue of research would be the inclusion a control group who merely repeats a task to control for this variable.

9.4 Pedagogical suggestions regarding use of TRTs

Three main findings of this study were: (i) that only a minority of language features pairs had revised themselves stimulated episodes of extended metatalk, (ii) that there were notable differences in degree of EWL between pairs during transcript revision, and these differences were especially marked when pairs examined the teacher's corrective feedback, and (iii) that differences in quality of metatalk and participation in making revisions were found only to have had a modest overall effect on retention of revisions. Given these findings, the 'elephant-in-the-room' question is whether I would recommend other teachers employ TRTs with their learners. The short answer is 'yes' but with certain cautions and caveats.

I recommend use of TRTs firstly because all eight students, regardless of their views on *pair*-revision of transcripts, viewed listening to and transcribing their oral performances as a valuable tool for self-assessment (section 8.1.2). All students pointed out in interviews that *self*-noticing when listening and transcribing was an important and eye-opening experience. Additionally, all eight students expressed the idea that, self-transcribing was 'more powerful' than only listening and that just listening to the recordings was 'insufficient' for noticing errors. All were supportive of the idea of being recorded and self-transcribing their oral production in the future. Secondly, while only four of the eight learners felt examining the teacher's revisions *with a partner* was necessary, all students wanted teacher feedback directed at their

transcripts. TRTs provide an efficient means of providing corrective-feedback (CF) directed at oral performance. Providing feedback *in-task* during students' oral task performance limits the teacher to providing CF only to those with whom they are interacting or observing at a given moment, and also risks interrupting the flow of student communication. TRTs are a means of providing *all* learners—not just those whose errors the teacher happens to overhear (and remember)—with corrective feedback.

All students, however, also pointed out that the transcript revision process needed streamlining. All learners noted that many of the mistakes contained in their *verbatim* transcripts could easily have been (and indeed were) corrected by themselves individually. Six students suggested that allowing them to 'clean up' their transcripts before submitting them to the teacher would streamline in-class transcript revision. I fully endorse this recommendation. If students revise their language whilst transcribing, this would streamline the in-class pair-revision process by reducing the number of 'obvious' errors a pair would need to go through and allow more time to focus on more problematic language features. If learners revise their transcripts before sending them to the teacher or bringing them to the classroom, then TRTs will have fostered independent study. If learners' self-revisions are incorrect, these inaccuracies would still be 'on the table' next lesson and available for scrutiny by the peer or teacher.

However, part of the process of 'cleaning up' transcripts the learners recommended included deletion of false starts, hesitations, and other dysfluencies. On the one hand, I see little loss in allowing these to be excised from transcripts before they are brought to class as dysfluencies never generated extended discussion and were often simply glossed over when transcripts were read aloud. I (the teacher) had also deleted learners' dysfluencies when revising transcripts. As previously noted, such dysfluencies were not considered in this study to be examples of

inaccurate language production and were deleted primarily to make it easier for learners to focus on my revisions of their *errors*. However, doing so may have given students the impression that a ‘quality’ live talk (speech) demands the absence of the pauses, false starts, and redundancies that typically attend and are a natural part of online delivery. Teachers need to be more explicit with learners than I was during this study that the presence of dysfluencies is both natural and unavoidable, even for the most competent L2 speakers.

A closely related caution is that there is a risk that, when making written corrections to render transcripts of oral language more accurate, learners and teachers misapply the grammatical standards of written discourse to spoken discourse. As Carter and McCarthy (2017) illustrate, it is now well recognized that the evidence from corpora of spoken discourse shows the standards of grammatical ‘correctness’ for written discourse differ markedly from that of spoken (especially conversational) grammar. However, Carter and McCarthy are arguably not discussing what Ellis and Barkhuizen (2005) would refer to as *absolute errors* which are linguistic forms that violate the TL ‘code’ and are not possible under TL rules, but rather noting that certain grammar forms and lexis that are common in spoken discourse are *dispreferred forms* (*ibid.*) in written discourse. In this study, all attempts were made to restrict restricted teacher-revision to absolute errors (see section 5.3.2), such as the following which would be ungrammatical regardless of modality: “So, poor people...amm...poor people almost difficult to be a doctor...so equality is not” (Yuta’s transcript; TRT 3). Most of the corpora of spoken English cited by Carter and McCarthy (2018)—such as the Nottingham Corpus—target informal and conversational language rather than more formal varieties, such as broadcast talk or public speaking. However, as the authors acknowledge in an earlier article, “there are many intermediate categories along a continuum from spoken to written” (McCarthy and Carter, 1995, p. 216). Before using TRTs to revise transcripts of learners’ public speaking, teachers should endeavour to ascertain where

monologic speeches fall within this continuum in terms of the grammar structures that are commonly employed. It would be a disservice to inadvertently teach learners to deliver stilted speeches that 'sound like a book' by misapplying revisions more appropriate to written compositions to online oral deliveries of (even formal) monologic speeches.

Although 6 out of 8 learners recommended being allowed to make revisions individually at home while transcribing their speeches, only 3 learners had recommended the NO PARTNER TRT plan when it was proposed during interviews. Of the remaining 5 learners, 4 (Naoto, Ken, Chika, and Momo) had unequivocally rejected this plan as they felt too much would be lost without having a partner to discuss revisions with; and the fifth learner, Aki, saying he would also prefer having a partner during editing with the added caveat that peer-editing is only useful if one's partner is as engaged by the process as oneself (Aki had felt this not to have been the case with his partner, Yuta).

Being able to employ and practice the revised language forms appeared important to the learners. In the student-teacher interviews, I proposed a potential alternative task cycle which dispensed with students redelivering their speeches. All eight learners roundly rejected this proposal noting that there would be far less or no incentive at all to revise their language output if the opportunity to use their revisions in repeat task performances were absent. Even though the extent to which learners discussed revisions (i.e., 'limited' vs. 'extended' metatalk) was found to have only a modest overall effect on students ability to retain revisions, the fact that revisions which had been the product of extended metatalk were more often retained was found to be statistically significant. Gains in ability of re-employ revisions, however modest, are gains nonetheless. Students like Asami and Naho (Pair 3), who produced the fewest episodes of extended metatalk when revising,

appear to be undermining their stated goal of subsequently employing revisions somewhat by not discussing revisions more extensively.

Therefore, although findings from a small-scale study such as this one cannot be generalized nor taken as predictive of the EWL of future learners, it is worth discussing how greater learner EWL might have been fostered during the revision process. As Svalberg (2009) rightly notes, EWL is very individual and context specific, and the views learners' expressed in interviews regarding the merits of TRTs were strongly colored by their having done them with a particular partner. It was, for example, impossible to determine if Naho and Asami's belief that Japanese students would only notice the same types of mistakes was a pre-existing belief or based solely on their experience of being each others' revision partners. These two learners were similar in terms of L2 proficiency, and their belief that Japanese peers have little of significance to the revision process might have been dispelled if either had been given the opportunity of working with a partner of higher proficiency such as Naoto or Chika.

Similarly, Naho and Asami had also evidenced largely passive uptake (i.e., without discussion or teacher consultation) of teacher-revisions, despite the majority of the revisions they had made themselves being equally accurate. Momo (Pair 2), in her interview, expressed that without the discussions with her partner (Chika) she also would not have had confidence in her own revisions and have 'just' adopted the teacher's. Therefore, Asami and Naho might similarly have discussed the possibility of their own revisions being as equally accurate as the teachers had they worked with other partners.

Aki and Yuta's EWL was impeded by a difference of opinion whether revising language for purely stylistic reasons was worthwhile: Aki felt it was and became increasingly frustrated that Yuta did not share this view. However, Naoto, Ken, and Chika were all inclined to making stylistic changes to transcripts. These learners

would have been accepting of Aki recommending such changes be made to their own transcripts. In short, changing the pairings of the learners likely would have ameliorated many of the impediments to EWL that appeared in this study.

A final impediment to EWL, expressed by Yuta, was the perception that, because the errors that appeared in his partner's delivery of a speech differed from those in his own, he could not make use of the corrections made to the partner's transcripts for his own benefit. In other words, Yuta had not found participating in the revision of his partner's transcript engaging because he felt he had no stake in helping others to revise transcripts of their individually delivered speeches. Yuta's partner, Aki, clearly expressed in interviews having felt frustrated by Yuta's lack of 'passion' when it had been his transcripts being revised. This impediment to EWL might have been avoided, or at least lessened, if the students had been asked to participate in *pair-team* debates (i.e., one student-pair debating another). If the team speeches had also been delivered by pairs, with each pair member delivering a given portion of the speech, Yuta might have felt he had some stake in the linguistic accuracy his partner's portion(s) of the speech.

In summary, as long as care is taken not to lead learners into misapplying the standards and expectations of written discourse when revising transcripts of oral production, I do recommend using TRTs in the L2 classroom. Use of TRTs need not be restricted to the monologic oral production of individual learners: Mennim (2003, 2007) had learners revise transcripts of group presentations, and Lynch (2001, 2007) had learners revising transcripts of pair role-plays. As suggested above, revision of learners' collaborative oral production may generate greater EWL than had the revision of the individual speeches in this study. Having students revise whilst transcribing would streamline the revision process. To further streamline the TRT process, teachers might consider having learners working together in class to simultaneously transcribe and revise recordings of their oral production. Resource-

wise, there are many devices today that can make voice recordings, from smart phones to laptops, which many students already use, in addition to many freely available software applications to help students to transcribe. For those with smaller sized classes, TRTs are a viable means of drawing learner attention to the accuracy of their oral output, and an efficient means of delivering teacher feedback to learners. The potential pedagogic applications of TRTs merit and deserve further study.

Appendices

Appendix A: Consent to Participate in Research (Students also provided with Japanese version)

Dear students

I am conducting research into language teaching and language learning for my doctoral studies at the University of Leicester. I would like to study the language that you use in this class in order to help me with my research.

Explanation of the Study

In this class, I am trying a new teaching method. My research will try to find out about improvements in your English if you are given the chance to repeat a task, especially if I also give you the chance to make your own corrections as well as being given corrections by myself as your teacher.

In this course you will be asked to perform English speaking tasks. Your performance of these speaking tasks will sometimes be audio recorded. You will then be asked to listen to the recording of what you said and to transcribe a portion of your speaking. You will have the chance to work with a partner to try to correct errors you notice in your transcript. When you are working with your partner to correct errors, what you say to each other will also be recorded. Then, additional help with errors in your transcripts will be provided by the teacher. You will then be given the chance to try performing the speaking task again, which will also be audio recorded.

The recordings of task performance, transcripts, and your discussions while making corrections are a regular part of the course, but I would also like to use them in my doctoral study.

Additionally, if you agree to do so, I would like to interview each student individually to hear your views and opinions on how the lessons may be improved, and include this information in my study. Findings from my doctoral study may be subsequently published in academic applied linguistics journals and/or presented at academic conferences.

Confidentiality

When I use your recordings, transcripts, and questionnaire/interview responses in my research, I will make sure that your identity remains *anonymous* and it will not be possible for anybody else to find out who said what. This means that I will not use your name. Your real name will never appear in my research writing.

Your participation

Giving your permission for me to use in my study recordings of what you say or write is voluntary. Your decision to participate—or not—will not affect your course grade. All students will be given equal consideration whether or not you agree to participate in this study. In the future, if you change your mind and no longer want to give permission, you can tell me and I will remove your recordings and writing from my study.

If you are happy for me to use your English for my research, please sign this piece of paper. If you have any questions or concerns, please ask me before signing this paper. I will give you a copy to keep.

Thank you
Jeremy Boston

CONSENT

Your name (Romanji): _____

I have read the information provided on this consent form. All my questions have been answered to my satisfaction. I voluntarily agree to allow Jeremy Boston to use my class recordings and writing for papers related to his research.

Signature:

Appendix B: Student profile questionnaire

1. Name
2. Age
3. Gender (male or female)
4. Degree program (major)
5. Why did you choose to major in that language?
6. Where were you born? If not Japan, how long have you been living in Japan?
7. What is your native language?
8. Have you spent any time staying in an English speaking country? If 'yes':
 - a. In what countries? (*example: Canada*)
 - b. How long? (*example: 2 months*)
 - c. At what age/how old were you? (*example: 14 years old*)
 - d. Why did you stay? (*example: home stay program*)
9. How long have you been studying English? ____ years?

☺Thank you for sharing this information ☺

Appendix C:

Recommended content-organization of debate speeches

(adapted from Lubetsky, Lebeau, and Harrington, 1999)

<u>Basic position speech organization</u>
- We are debating the resolution [<i>e.g., Smoking should be banned</i>]. - I agree/disagree with this resolution for [<i>four</i>] reasons: [<i>health</i>], [<i>hospital costs</i>], [<i>pollution</i>], and [<i>that it's unattractive</i>]. - My first point is [<i>health</i>]: [<i>Explain and support this reason</i>]. - My second point is _____: _____. - My third point... - My fourth point.... - In conclusion, I have talked about [<i>health</i>], [<i>hospital costs</i>], [<i>pollution</i>], and [<i>that it's unattractive</i>]. - I have clearly shown that <u><i>Smoking should be banned</i></u>.

Appendix D: Example of unrevised student-generated transcript

Today's resolution is all schools should be public: private schools should be abolished. I support this resolution. I have three reasons. Aa,,,,, equality of education, diversity, and quality of education.

My first reason is quality of education. Aaa,,first,,ahmm basically, education must not have a difference amm,, between rich people and poor people. And,,, we must ,,must have opportunity to have a,,,,,education to the rich people and poor people. Amm,,,, we have the right to receive education amm,,,, whether the people who have aaa,,, rich parent or not. Amm,,, for example, ah,, if ,, if you want to be a doctor, your parent have to be a rich. So, poor people,,, amm,, poor people almost difficult to be a doctor. So, this is not equal. So,, amm,, all school should be public.

And,,, my second reason is diversity. Ah,,, ummm,, public school have a diversity. Umm for example public school have many kind of people amm,,, comes from different,, umm different places or different background and different parents. So, many kind of people are there. So you can get rich humanity to,,,, umm,, when you meet many kind of people, you can get ,,you can get rich humanity. And good command to communicate with the many kind of people through your school life. It is the feature of ,,,,amm most good point of umm,, public school.

Amm,,, my third reason is quality of education. Amm,, basically, ah,,, in Japan, there are a lot of schools, ,,and,, and,, private school, ahh ,, not every private school have a high quality because of its amm,, its tuition its high tuition or high fee. And,, there are good public school ahh,, umm,, there are good high quality education schools too. Amm,, in my experience, ah,, my junior high school friends go to a publi,, amm many friends go to private school but many ,, but also there are a lot of people amm,,, not educated. So,,, I think it is the waste of money, and waste of time. So, private school should be abolished.

Amm,, I have talked about equality of education, diversity, quality of education. I have clearly shown that all schools should be public: private schools should be abolished.

Appendix E: End-of-course interview guide

THIS YEAR'S PLAN

<u>1 Lesson</u> SAY YOUR SPEECH THE FIRST TIME		
	<i>Homework</i> TRANSCRIBE 1 st SPEECH	
<u>2 Lesson</u> EDIT TRANSCRIPTS WITH PARTNER		
	<i>Homework</i> NONE	
<u>3 Lesson</u> SAY YOUR SPEECH THE SECOND TIME + DEBATE IDEAS		
	<i>Homework</i> TRANSCRIBE 2 nd SPEECH	

QUESTIONS

Listening to and Transcribing Recordings

1. What did you think about listening to recordings of your speech?
 - a. How did you feel when you listened to your own performance?
 - b. Was there anything you found useful about listening to recordings of yourself speaking?
2. What did you think about transcribing recordings of your speech?
 - a. Was there anything you found useful about this homework?

- b. How long did it take you to transcribe a speech?
- c. Was anything difficult about this homework?
- d. What do you think about being recorded and transcribing in future English courses?

Working with a Partner

- 3. What did you think about improving speeches with a partner?
 - a. What did you like about editing with a partner?
 - b. What did you NOT like about editing with a partner?
 - c. When you improve your speech, some students are not sure it is useful to have a partner. What do you think about this?
 - d. Some students are not sure it is useful for you to help the other student's to improve the other students' speech. What do you think about this?
 - e. When looking at the teacher's corrections, some students are not sure it is useful having a partner to talk to. What do you think?

Using English Only

- 4. How do you feel about improving speeches while speaking only in English?
 - a. My Boss felt discussing the English language in Japanese was not suitable for an English speaking and listening course. What do you think?
 - b. Next year, I could allow students to speak Japanese when improving speeches. What do think about this idea?

Advice for Teacher

5. Here is a new plan for next year. What do you think about this plan? Do you recommend I try this?

PLAN 1: NO PARTNER

Lesson 1 DELIVER SPEECH FIRST TIME	
Homework (i)TRANSCRIBE SPEECH + (ii)MAKE IMPROVEMENTS <u>BY YOURSELF</u> AND SEND TO TEACHER + (iii)GET TEACHER FEEDBACK	
Lesson 2 DELIVER SPEECH SECOND TIME + DEBATE	

6. Here is another new lesson plan for next year. What do you think about this plan? Do you recommend I try this?

PLAN 2: NO SECOND SPEECH

Lesson 1 DELIVER SPEECH <u>and</u> DEBATE	
	Homework TRANSCRIBE SPEECH and SEND TO TEACHER
Lesson 2 IMPROVE SPEECH <u>WITH PARTNER</u> + GET TEACHER FEEDBACK	

7. If you had to choose only one plan: (1) *THIS YEAR'S PLAN*; (2) *NO PARTNER PLAN*; or (3) *NO SECOND SPEECH PLAN*, which would you choose? Why?
8. Do you have any extra advice to help me improve the course next year?

Appendix F: Transcript conventions for student interaction data

ALL CAPS	Students reading verbatim what is written on their transcriptions, or extracts from the transcripts to indicate original, unrevised output.
'Single quotes'	Isolates the forms the students are recommending or discussing.
.	Falling "Final" intonation.
?	Question intonation, not necessarily at end of sentence/utterance.
! <u>Emphasis</u>	Shows student excitement or surprise. Students consciously stress/emphasize word or part of word to make it clear to revision partner.
w-o-r-d	Spelling out of word/words.
[	Onset of overlapping speech.
=	Latching or no gap between turns.
XXX	Unintelligible word/words.
...	Short pause of 1-3 seconds.
<5>	Longer pause. Number indicates time in seconds <5> and minutes <5:00>.
<i>Italics</i>	Japanese utterance.
(<i>Italics</i>)	Translation of Japanese utterance.
(Comment/other details)	Transcriber's commentary.

Appendix G: Full and completed LRE analysis table

Figure 5.5: Matched sentences: Initial, teacher, and final revisions

Original [Line 6]	SO WE WILL BE ABLE TO HAVE MANY MONEY				
Initial Revision [Line 6]	SO WE WILL BE ABLE TO HAVE <u>MUCH</u> MONEY				
Teacher Revision [Line 6]	SO WE WILL BE ABLE TO HAVE <u>A LOT OF</u> MONEY				
Final Revision [Line 6]	SO WE WILL BE ABLE TO HAVE (1) <u>A LOT OF</u> MONEY				
Pair-talk: Initial revision (A = Asami [author]; N = Naho [revision partner])					
26	A: SO WE WILL BE ABLE TO HAVE MANY? MONEY.				LRE 1
27	N: ‘Much’ money.				
28	A: ‘much’.				
LRE Analysis (Initial revision)					
LRE	Focus	Outcome	Quality of Metatalk	Source Initiation	Source Resolution
1	L-LRE (word choice ‘much’)	-/+	L	SI	OR
LRE Analysis (Final revision)					
41	A: SO WE WILL BE ABLE TO HAVE A LOT OF MONEY...oh, ‘a lot’...				LRE 1
42	N: Much money is change to ‘a lot’. Mm (got it).				
LRE Analysis (Final revision)					
LRE	Focus	Prior Noticing	Outcome	Source Revision	Teacher Consulted
1	L-LRE (word choice)	Y	+/+	T	N

References

- Ahmadian, M.J. (2011) 'The effect of 'massed' task repetitions on complexity, accuracy and fluency', *The Language Learning Journal*, 39(3) pp. 269-280.
- Al-Hejin, B. (2004) 'Attention and awareness: Evidence from cognitive and second language research', *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* 4 (1), pp. 1-22.
- Breen, M.P. (1987) 'Learner contribution to the task design', in Candlin, C.N. and Murphy, D. (eds.), *Language learning tasks*, Saddle River: Prentice-Hall International, pp. 23-46.
- Brooks, L., and Swain, M. (2009) 'Languaging in collaborative writing: Creation of and response to expertise', in Mackey, A. and Polio, C. (eds.), *Multiple perspectives on interaction in SLA*, Mahwah, NJ: Lawrence Erlbaum, pp. 58-89.
- Burns, A. (2005) 'Action research: an evolving paradigm?', *Language Teaching*, 38 (2), pp. 57-74.
- Bygate, M. (1996) 'Effects of task repetition: Appraising the developing language of learners,' in Willis, J. and Willis, D. (eds.), *Challenge and change in language teaching*, Oxford: Heinemann, pp. 136-146.
- Bygate, M. (2001) 'Effects of task repetition on the structure and control of oral language', in Bygate, M., Skehan, P., and Swain, M. (eds.), *Researching Pedagogic Tasks, Second Language Learning, Teaching and Testing*, Harlow: Longman, pp. 23-48.
- Bygate, M., and Samuda, V. (2005) 'Integrative planning through the use of task repetition', in Ellis, R. (ed.), *Planning and task performance in a second language*, Amsterdam: John Benjamins, pp. 37-74.
- Carter, R., and McCarthy, M. (2017) 'Spoken Grammar: Where are we and where are we going?', *Applied Linguistics*, 38 (1), pp. 1-20.
- Cohen, L., and Manion, L. (1994) *Research Methods in Education* (4th ed.) London, UK: Routledge.
- Cohen, L., Manion, L., and Morrison, K. (2000) *Research Methods in Education*, (5th ed.) London: Routledge.

- Corder, S. P. (1974) 'Error analysis', in Allen, J. and Corder, S. (eds.), *The Edinburgh course in applied linguistics: Techniques in applied linguistics*, Oxford, UK: Oxford University Press
- Coughlan, P. and Duff, P. A. (1994) 'Same task, different activities: Analysis of SLA task from an activity theory perspective', in Lantolf, J.P. and Appel, G., (eds.) *Vygotskian approaches to second language research*, Norwood, NJ: Ablex, pp. 173-193.
- Cresswell, J. (1998) *Qualitative inquiry and research design: Choosing among five traditions*, Thousand Oaks, CA: Sage.
- Donato, R. (1994) 'Collective scaffolding in second language learning', in Lantolf, J.P. and Appel, G. (eds.), *Vygotskian approaches to second language research*, Norwood, NJ: Ablex, pp. 33–56.
- Dörnyei, Z. (2009) *The psychology of second language acquisition*, Oxford, UK: Oxford University Press.
- Duff, P. (2008) *Case Study Research in Applied Linguistics*, New York, NY: Routledge.
- Eckerth, J. (2008) 'Investigating consciousness raising tasks: Pedagogically targeted and non-targeted learning gains', *International Journal of Applied Linguistics*, 18(2), pp. 119–145.
- Ellis, R. (1994) *The Study of Second Language Acquisition*, Oxford, UK: Oxford University Press.
- Ellis, R. (1997) *SLA Research and Language Teaching*, Oxford: Oxford University Press.
- Ellis, R. (2003) *Task-based language learning and teaching*, Oxford, UK: Oxford University Press.
- Ellis, R. (2008) *The study of second language acquisition*, 2nd edn., Oxford, UK: Oxford University Press.
- Ellis, R. (2009) 'Implicit and explicit learning, knowledge and instruction', in Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J. and Reinders H., *Implicit and explicit knowledge in second language learning, testing and teaching*, Bristol, UK: Multilingual Matters, pp. 3-25.
- Ellis, R. (2015) *Understanding second language acquisition*, 2nd Edn., Oxford, UK: Oxford University Press.

- Ellis, R., and Barkuizen, G. (2005) *Analysing learner language*, Oxford, UK: Oxford University Press.
- Ellis, R., Basturkmen, H., and Loewen, S. (2001) 'Learner uptake in communicative ESL lessons', *Language Learning*, 51(2), pp. 281–318.
- Erlam, R. (2009) 'The elicited oral imitation test as a measure of implicit knowledge', in Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J. and Reinders H., *Implicit and explicit knowledge in second language learning, testing and teaching*, Bristol, UK: Multilingual Matters, pp. 65-93.
- Fortune, A., and 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), pp. 143–60.
- Foster, P. (1998) 'A classroom perspective on the negotiation of meaning', *Applied Linguistics*, 19(1), pp. 1-23.
- Gass, S., Mackey, A., Fernandez, M., and Alvarez-Torres, M. (1999) 'The effects of task repetition on linguistic output', *Language Learning* 49 (4), pp. 549–80.
- Holec, H. (1981) *Autonomy and foreign language learning*, Oxford, UK: Pergamon.
- Hyland, F. (1998) 'The impact of teacher written feedback on individual writers', *Journal of Second Language Writing*, 7 (3), pp. 255–286.
- Kim, Y. J. (2008) 'The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary', *Modern Language Journal*, 92(1), pp. 114-130.
- Kim, Y. J., and McDonough, K. (2008) 'The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners', *Language Teaching Research*, 12(2), pp. 211-234.
- Kim, Y. J., and McDonough, K. (2011) 'Using pretask modeling to encourage collaborative learning opportunities', *Language Teaching Research*, 15(3), pp. 183-199.
- Knouzi, I., Swain, M., Lapkin, S., and Brooks, L. (2010) 'Self-scaffolding mediated by languaging: Microgenetic analysis of high and low performers', *International Journal of Applied Linguistics*, 20(1), pp. 23–49.
- Kowal, M. and Swain, M. (1994) 'Using collaborative language production tasks to promote students' language awareness', *Language Awareness*, 3(1), pp. 73–93.

- Kouzlin, A. (1986) 'Vygotsky in context', in Kouzlin, A. (ed.), *L. S. Vygotsky: Thought and language* (revised edn.). Cambridge, MA: Cambridge University Press, pp. xi-lvi.
- Lantolf, J. P. (2000a) 'Introducing sociocultural theory', in J. P. Lantolf, J.P. (ed.), *Sociocultural theory and second language learning*, Oxford: Oxford University Press, pp. 1–26.
- Lantolf, J. P. (2000b) 'Second language learning as a mediated process', *Language Teaching*, 33 (1), pp. 79–96.
- Lantolf, J. P. (2005) 'Sociocultural and second language learning research: an exegesis' in Hinkel, E. (ed.), *Handbook of Research on Second Language Teaching and Learning*, Mahway, NJ: Lawrence Erlbaum, pp. 335-354.
- Lantolf, J. P. (2006) 'Sociocultural theory and L2 development: State-of-the-art', *Studies in Second Language Acquisition*, 28, pp. 67–109.
- Lantolf, J. P., and Thorne, S. L. (2007) 'Sociocultural theory and second language learning', in Van Pattern, B., and Williams, J. (eds.), *Theories in second language acquisition*, Mahwah, NJ: Lawrence Erlbaum, pp. 201-224.
- Lapkin, S., Swain, M., and Smith, M. (2002) 'Reformulation and the learning of French pronominal verbs in a Canadian French immersion context', *Modern Language Journal*, 86 (4), pp. 485–507.
- Larsen-Freeman, D. (1997) 'Chaos/complexity science and second language acquisition', *Applied Linguistics*, 18 (2), pp. 141-165.
- Lee, I. (2007) 'Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning?', *Assessing Writing*, 12 (3), pp. 180–198.
- Leeser, M. J. (2004) 'Learner proficiency and focus on form during collaborative dialogue', *Language Teaching Research*, 8 (1), pp. 55-81.
- Leow, R.P. (1997) 'Attention, awareness, and foreign language learning', *Language Learning*, 47 (3), pp. 467-505.
- Leow, R. P. (2000) 'A study of the role of awareness in foreign language behavior: Aware vs. unaware learners', *Studies in Second Language Acquisition*, 22 (4), pp. 557-584.
- Leow, R. P. (2001) 'Attention, awareness and foreign language behavior. *Language Learning*, 51, pp. 113–155.

- Little, D. (2007) 'Language learner autonomy: Some fundamental considerations revisited', *Innovation in Language Learning and Teaching*, 1 (1), pp. 14–29.
- Loewen, S. (2005) 'Incidental focus on form and second language learning', *Studies in Second Language Acquisition*, 27 (3), pp. 361-386.
- Loewen, S. (2009) 'Grammaticality judgement tests and the measure of implicit and explicit L2 knowledge' in Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J. and Reinders H., *Implicit and explicit knowledge in second language learning, testing and teaching*, Bristol: Multilingual Matters, pp. 94-112.
- Long, M. (1996) 'The role of the linguistic environment in second language acquisition', in Ritchie, W. C., and Bahtia, T. K. (eds.), *Handbook of second language acquisition*, New York: Academic Press, pp. 413-68.
- Long, M. (2006) 'Recasts in SLA: The story so far', in M. Long (ed.), *Problems in SLA*, Mahwah, NJ: Erlbaum, pp. 75-116.
- Loschky, L., and Bley-Vroman, R. (1993) 'Grammar and task-based methodology,' in Crookes, G. and Gass, S. (1993) *Tasks and language learning: Integrating theory and practice*. Clevedon: Multilingual Matters, pp. 123-167.
- Lubetsky, M., Lebeau, C. and Harrington, D. (1999) *Discover debate: Basic skills for supporting and refuting opinions*, Medford, Oregon: Language Solutions Inc.
- Lynch, T. (2001) 'Seeing what they meant: transcribing as a route to noticing', *ELT Journal*, 55 (2), pp.124–32.
- Lynch, T. (2007) 'Learning from the transcripts of an oral communication task', *ELT Journal*, 61(4), pp. 311–320.
- Lyster, R. and Ranta, L. (1997), 'Corrective feedback and learner uptake: Negotiation of form in communicative classrooms', *Studies in Second Language Acquisition*, 19 (1), pp. 37-66.
- Mackey, A., Gass, S., and McDonough, K. (2000) 'Do learners perceive interactional feedback', *Studies in Second Language Acquisition*, 22 (3), pp. 471-497.
- McCarthy, M., and Carter, R. (1995) 'Spoken grammar: what is it and how can we teach it?', *ELT Journal*, 49 (3), pp. 207–218.
- Malmqvist, A. (2005) 'How does group discussion in reconstruction tasks affect written language output?', *Language Awareness*, 14 (2), pp. 128-141.

- McDonough, K., and Sunitham, W. (2009) 'Collaborative dialogue between Thai EFL learners during self-access computer activities', *TESOL Quarterly*, 43(2), pp. 231-254.
- Mennim, P. (2003) 'Rehearsed oral L2 output and reactive focus on form', *ELT Journal* 57 (2), pp. 130–138.
- Mennim, P. (2007) 'Long term effects of noticing on oral output', *Language Teaching Research*, 11 (3), pp. 265–280.
- Mennim, P. (2012) 'Learner negotiation of L2 form in transcription exercises', *ELT Journal*, 66 (1), pp. 52-61.
- Nagatomo, D.H. (2012) *Exploring Japanese university English teachers' professional identity*, Bristol, UK: Multilingual Matters.
- Ohta, A. S. (2000) 'Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar', in Lantolf, J. P. (ed.), *Sociocultural theory and second language learning*, Oxford, UK: Oxford University Press, pp. 51-78.
- Ohta, A. S. (2001) *Second language acquisition processes in the classroom: Learning Japanese*, Mahwah, NJ: Lawrence Erlbaum.
- Paulus, T. M. (1999) 'The effect of peer and teacher feedback on student writing', *Journal of Second Language Writing*, 8 (3), pp. 265–289.
- Poehner, M. E. (2008) 'Both sides of the conversation: The interplay between mediation and learner reciprocity in dynamic assessment', in Lantolf, J. P. and Poehner, M. E (eds.), *Sociocultural theory and the teaching of second languages*, London: Equinox, pp. 33–56.
- Qi, D.S. and Lapkin, S. (2001) 'Exploring the role of noticing in a three-stage secondlanguage writing task', *Journal of Second Language Writing*, 10 (4), pp. 277–303.
- Reinders, H. (2010) *The effects of task type and instructions on second language acquisition*, Newcastle upon Tyne: Cambridge Scholars Press.
- Robinson, P. (1995) 'Attention, memory and the "noticing" hypothesis', *Language Learning*, 45 (2), pp. 283–331.

- Robinson, P. (2003), 'Attention and memory during SLA', in Catherine J., Doughty, C., and Long, M. (eds.) *The handbook of second language acquisition*, New York: Blackwell Publishing, pp. 631-678.
- Robinson, P. (2015), 'Attention and Awareness', in Cenoz, J., Gorter, D., and May, S. (eds.) *Language awareness and multilingualism: Encyclopedia of language and education*, 3rd edn., New York: Springer, pp. 1-10.
- Rosa, E. M. and Leow, R. P. (2004) 'Awareness, different learning conditions, and second language development', *Applied Psycholinguistic*, 25, pp. 269-292.
- Rosa, E., and O'Neill, M. D. (1999) 'Explicitness, intake, and the issue of awareness' *Studies in Second Language Acquisition*, 21 (3), pp. 511-556.
- Sachs, R., and Polio, C. (2007) 'Learners' uses of two types of written feedback on a L2 writing revision task', *Studies in Second Language Acquisition*, 29 (1), pp. 67-100.
- Sato, M. (2017) 'Interaction mindsets, interactional behaviors, and L2 Development: An affective-social-cognitive model', *Language Learning*, 67 (2), pp. 249-283.
- Schmidt, R., (1990) 'The role of consciousness in second language learning', *Applied linguistics*, 11(2), pp.129-158.
- Schmidt, R., (1995) 'Consciousness and foreign language learning: A tutorial on the role of attention and awareness', in Schmidt, R. (ed.) *Attention and awareness in foreign language learning*. Honolulu: University of Hawaii, pp. 1-63.
- Schmidt, R. (2001) 'Attention', in Robinson, P. (ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press, pp. 3-32.
- Schmidt, R. (2010) 'Attention, awareness, and individual differences in language learning', in Chan, W. M., Chi S., Cin, K., Istanto, N. J., Nagami, M.J., Sew, W., Suthiwan, T., and Walker, I. (eds.) *Proceedings of CLaSIC 2010*. Singapore: National University of Singapore, Centre for Language Studies, pp. 721-737.
- Schmidt, R. and Frota, S., (1986) 'Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese', in Day, R. (ed.) *Talking to learn: Conversation in second language acquisition*. Rowley, MA: Newbury House, pp.237-326.

- Segalowitz, N. and Lightbown, P. (1999) 'Psycholinguistic approaches to SLA', *Annual Review of Applied Linguistics*, 19, pp. 43–63.
- Sharwood Smith, M. (1993) 'Input enhancement in instructed SLA: Theoretical bases', *Studies in Second Language Acquisition*, 15, pp. 165–179.
- Skehan, P. (1996) 'Second language acquisition research and task-based instruction', in Willis, J. and Willis, D (eds.) *Challenge and change in language teaching*, Oxford: Oxford University Press, pp. 17-30.
- Skehan, P. (1998) *A cognitive approach to language learning*, Oxford: Oxford University Press.
- Skehan, P., Bei, X., Li, Q., and Wang, Z. (2012) 'The task is not enough: Processing approaches to task-based performance', *Language Teaching Research*, 16(2), pp. 170–187.
- Stillwell, C., B., Curabba, K., Alexander, A. Kidd, E. Kim, P. Stone, and Wyle, C. (2010) 'Students transcribing tasks: noticing fluency, accuracy, and complexity', *ELT Journal* 64 (4), pp. 445–455.
- Stone, P. (2015) 'Video reflexive ethnography as a tool for better classroom practice' *The Language Teacher*, 39 (6), pp. 10-14.
- Storch, N. (1997) 'The editing task of adult ESL learners', *Language Awareness*, 6 (4), pp. 221– 232.
- Storch, N. (2002) 'Patterns of interaction in ESL pair work', *Language Learning*, 52 (1), pp. 119–158.
- Storch, N. (2004) 'Using activity theory to explain differences in patterns of dyadic interactions in an ESL class', *The Canadian Modern Language Review*, 60(4), pp. 457– 480.
- Storch, N. (2007) 'Investigating the merits of pair work on a text editing task in ESL classes', *Language Teaching Research* 11 (2), pp. 143– 161.
- Storch, N. (2008) 'Metatalk in pair work activity: Levels of engagement and implications for language development', *Language Awareness*, 17(2), pp.95-114.
- Storch, N. (2013) *Collaborative writing in L2 classrooms*, Bristol, England: Multilingual Matters.
- Storch, N., and Aldosari, A. (2013) 'Pairing learners in pair work activity', *Language Teaching Research*, 17 (1), pp. 31– 48.

- Storch, N., and Wigglesworth, G. (2010) 'Learners' processing, uptake, and retention of corrective feedback on writing', *Studies in Second Language Acquisition*, 32, pp. 303-334.
- Strauss, A., and Corbin, J. (1998) *Basics of qualitative research*, London: Sage publications.
- Svalberg, A.M.L. (2007) 'Language awareness and language learning', *Language Teaching*, 40 (4), pp. 287-308.
- Svalberg, A.M.L. (2009) 'Engagement with language: Interrogating a construct', *Language Awareness*, 18 (3), pp.242-258.
- Svalberg, A.M.L (2012) 'Language Awareness in language learning and teaching: A research agenda', *Language Teaching*, 45 (3), pp. 376- 388
- Swain, M. (1993) 'The output hypothesis: Just speaking and writing aren't enough', *The Canadian Modern Language Review*, 50 (1), pp. 158-164.
- Swain, M. (1995) 'Three functions of output in second language learning', in Cook, G.and Seidlhofer, B. (eds.), *Principles and practice in applied linguistics: Studies in honour of H.G. Widdowson*, Oxford: Oxford University Press, pp. 125-144.
- Swain, M. (1998) 'Focus on form through conscious reflection', in Doughty, C. and Williams, J. (eds.), *Focus on form in classroom second language acquisition*, Cambridge: Cambridge University Press, pp. 64-81.
- Swain, M. (2000) 'The output hypothesis and beyond: Mediating acquisition through collaborative dialogue', in Lantolf, J. (ed.), *Sociocultural theory and second language learning*, Oxford, UK: Oxford University Press, pp. 97-114.
- Swain, M. (2001) 'Examining dialogue: another approach to content specification and to validating inferences drawn from test scores', *Language Testing*, 18 (3), pp. 275-302.
- Swain, M. (2005) 'The output hypothesis: Theory and research', in Hinkel, E. (ed.), *Handbook of research in second language teaching and learning*, Mahwah, NJ: Lawrence Erlbaum, pp. 471-483
- Swain, M. (2006) 'Languaging, agency and collaboration in advanced language proficiency', in Byrnes, H. (ed.), *Advanced language learning: The contribution of Halliday and Vygotsky*, London, UK: Continuum, pp. 95-108.

- Swain, M. (2010) 'Talking-it-through: Linguaging as a source of learning', in Batstone, R. (ed.), *Sociocognitive perspectives on language use/learning*, Oxford, UK: Oxford University Press, pp. 112–130.
- Swain, M. (2013) 'The inseparability of cognition and emotion in second language learning', *Language Teaching*, 46 (2), pp. 195-207.
- Swain, M., Brooks, L., and Tocalli-Beller, A. (2002) 'Peer-peer dialogue as a means of second language learning', *Annual Review of Applied Linguistics*, 22, pp. 171–85.
- Swain, M., and Lapkin, S. (1995) 'Problems in output and the cognitive processes they generate: A step towards second language learning', *Applied Linguistics*, 16 (3), pp. 371-391.
- Swain, M., and Lapkin, S. (1998) 'Interaction and second language learning: Two adolescent French immersion students working together', *Modern Language Journal*, 82 (3), pp. 320–37.
- Swain, M., and Lapkin, S. (2001) 'Focus on form through collaborative dialogue: Exploring task effects', in Bygate, M., Skehan, P., and Swain, M. (eds.), *Researching pedagogic tasks: Second language learning, teaching and testing*, London, UK: Longman, pp. 99–118.
- Swain, M., and Lapkin, S. (2002) 'Talking it through: Two French immersion learners' response to reformulation', *International Journal of Educational Research*, 37, pp. 285– 304.
- Swain, M., Lapkin, S., Knouzi, I., Suzuki, W. and Brooks, L. (2009) 'Linguaging: University students learn the grammatical concept of voice in French', *The Modern Language Journal* 93, pp. 5– 29.
- Tocalli-Beller, A., and Swain, M. (2005) 'Reformulation: The cognitive conflict and L2 learning it generates', *International Journal of Applied Linguistics*, 15 pp. 5–28.
- Tomlin, R. S., and Villa, V. (1994) 'Attention in cognitive science and second language acquisition', *Studies in Second Language Acquisition*, 16 (2), pp. 183–203.
- VanPatten, B. (1990) 'Attending to form and content in the input: An experiment in consciousness', *Studies in second language acquisition*, 12(03), pp. 287-301.

- Villamil, O.S., and De Guerrero, M. (1996) 'Peer revision in the L2 Classroom: Social- cognitive activities, mediating strategies, and aspects of social behavior', *Journal of Second Language Writing*, 5(1), pp. 51-57.
- Vygotsky, L. S. (1978) *Mind in society*, Cambridge, MA: MIT Press.
- Watanabe, Y. (2008) 'Peer-peer interaction between L2 learners of different proficiency levels: Their interactions and reflections', *Canadian Modern Language Review*, 64 (4), pp. 605–636.
- Watanabe, Y. and Swain, M. (2007) 'Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners', *Language teaching research*, 11(2), pp.121-142.
- Williams, J. (1999) 'Learner-generated attention to form', *Language Learning*, 49(4), pp. 583-625.
- Williams, J. (2001) 'The effectiveness of spontaneous attention to form. *System*, 29(3), pp. 325-340.
- Williams, J. N. (2005), 'Learning without awareness', *Studies in Second Language Acquisition*, 27 (2), pp. 269-304.
- Wong, W. (2001) 'Modality and attention to meaning and form in the input', *Studies in Second Language Acquisition*, 23, pp. 345-368.
- Wragg, T. (2002) 'Interviewing', in Coleman, M. and Briggs, A.R.J. (eds.), *Research methods in educational leadership and management*, London, UK: Paul Chapman Publishing, pp. 143-158.
- Yang, M., Badger, R., and Yu, Z. (2006) 'A comparative study of peer and teacher feedback in a Chinese EFL writing class', *Journal of Second Language Writing*, 15 (3), pp. 179–200.
- Yin, R. (2003) *Case study research: Design and methods* (3rd ed.), Thousand Oaks, CA: Sage.
- Zhao, H. (2010) 'Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom', *Assessing Writing*, 15 (1), pp. 3-17.

