Transcription as an Initiator of Noticing Language Form and Use: 
an EFL Longitudinal Study

by

Margaret M. C. So

A thesis submitted in conformity with the requirements 
for the degree of Doctor of Philosophy
Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Margaret M.C. So (2015)
ABSTRACT

When second language learners speak, do they know if they are making mistakes? And what would they do if they knew? This research investigates noticing in relation to speaking. Noticing is an important first step in the process of second language learning. Many research studies have shown that different instructional approaches can draw learner attention to particular forms so that they can become noticed. However, why a learner chooses to attend to certain input and not others remains little understood.

In this longitudinal study, transcription was used as a means to initiate noticing. Twenty-four Hong Kong ESL senior secondary school students transcribed the oral presentations they had done as part of a mock exam practice. With the aid of the transcripts, they reflected on their oral production with a classmate or on their own, after which they were twice given time to prepare a revision and then repeat their performance. Stimulated recalls provided additional opportunities for noticing. Learners underwent the procedure three times over the course of one school semester.

I found that participants were able to discover their own grammar and pronunciation errors through their transcripts, and were able to correct most of their mistakes at the next opportunity. The main difference was in how collaborative pairs evolved in their process of noticing over time. The study found that the task itself had an impact on how the participants
prioritized and acted on what they noticed. I conclude that transcription is a useful activity that
leads students to notice their own speech production, while collaborative dialogue should have a
significant place in the language classroom as shown in the qualitative, but not the quantitative,
data.
ACKNOWLEDGEMENTS

This dissertation exists because of the generous support and encouragement from so many. Firstly, I wish to thank my thesis supervisor, Dr. Merrill Swain. To her I owe my deepest gratitude. Her inspirational scholarship opened new avenues of thinking for me, while her guidance, unparalleled patience and encouragement sustained my efforts through this long journey. I also wish to express my heartfelt thanks to my committee members, Dr. Sharon Lapkin and Dr. Nina Spada, for their patience and invaluable comments on my work. As well, I wish to thank Dr. Hossein Nassaji for his insightful comments and questions in his role as external examiner.

I would also like to thank my colleagues, in particular: Khaled Barkoui, Tae-Young Kim, Colette Peters, Wataru Suzuki, and Antonella Valeo. Their willingness to share their knowledge and humour my questions fostered a community which I was privileged to have been a part of.

Special thanks are due to my Hong Kong participants who gave of their valuable time. I am also deeply grateful to Petunia Kingsley for her support of my endeavours while I was in Hong Kong. As well, I wish to thank Monique Lok for lending her expertise in this study. In Canada, I wish to thank my friends, who have encouraged me and facilitated my studies, in particular, Clodagh Slowey.

I am grateful for the financial support I received in the form of an Ontario Graduate Scholarship, and a School of Graduate Studies Travel Grant, which facilitated this research.

Lastly, I wish to express my gratitude to my family. I thank my Hong Kong relatives, who nourished me with soup and challenged me to long hikes in between data collection. My family’s unfailing support and encouragement through the years has been a source of strength. I thank my parents for giving me wings to fly, and teaching me courage; Rebecca and Paul for their unstinting support; Dominic, Miang Boon, Jacob and Mikayla: for their wisdom and stories,
which kept me grounded. I also thank my Canadian family: Don and Joyce; Susan and Don; Margaret and Brian; for their encouragement, forbearance, and unwavering support through the years, particularly when childcare was needed. Closer to home, I thank Luke T-S and O-S Hart for the joy they bring. Finally, I thank Andrew for encouraging me to start on this road, and finishing this very long journey with me. Thank you for your patience, understanding and love.
TABLE OF CONTENTS

ABSTRACT ............................................................................................................................... ii
ACKNOWLEDGEMENTS .......................................................................................................... iv
TABLE OF CONTENTS ............................................................................................................ vi
LIST OF TABLES ....................................................................................................................... ix
LIST OF FIGURES ................................................................................................................... x
LIST OF APPENDICES .......................................................................................................... xi

CHAPTER 1: INTRODUCTION ................................................................................................ 1
  Rationale .............................................................................................................................. 5
  Organization of thesis ........................................................................................................ 6

CHAPTER 2: LITERATURE REVIEW ..................................................................................... 8
  Section One: Noticing ........................................................................................................ 8
    Attention in SLA research .............................................................................................. 10
  Section Two: Collaborative Dialogue ............................................................................ 16
  Section Three: Transcription Studies ............................................................................ 20
  Conclusion ......................................................................................................................... 32
  Research Questions ......................................................................................................... 33

CHAPTER 3: METHODOLOGY ............................................................................................... 34
  Section One: Background ............................................................................................... 34
    Context of the Study ....................................................................................................... 34
    Participants ....................................................................................................................... 35
    Use of English: Paper D ................................................................................................. 36
  Section Two: Research Design ....................................................................................... 38
    Rationale for using Paper D ......................................................................................... 38
    Pilot study ....................................................................................................................... 39
    How the Research Questions and Instructional Intervention Evolved ....................... 39
    Organization of Participants ......................................................................................... 40
    Data Collection Procedure: the Sessions .................................................................. 42
  Section Three: Data treatment ....................................................................................... 46
    Organizing the Data ....................................................................................................... 46
    Coding for Engagement ............................................................................................... 53
    Principles Used for Reporting Data ............................................................................. 54
    Limitations of Quantitative Data ................................................................................ 54

CHAPTER 4: THE QUANTITATIVE DATA AND RESULTS .................................................. 56
  Section One ....................................................................................................................... 56
    Grammar ......................................................................................................................... 57
    Pronunciation ................................................................................................................ 59
    Content ........................................................................................................................... 59
  Section Two ....................................................................................................................... 61
    Point of Noticing .......................................................................................................... 61
    Judgement Possible (JP) and Trackable items ............................................................ 61
    RQ 2b: When given opportunities to repeat their speech performance, what do learners do with the items they have noticed? ......................................................... 66
LIST OF TABLES

Table 2 A Summary of the Data Collection Procedure Cycle for One Session ................................................................. 44
Table 3 An Example of a Coded Item From a Pair Session .................................................................................................................. 48
Table 4 An Explanation of What was Entered in Each Column of the Coding Grid ................................................................. 49
Table 5 An example of First Noticing and Product Corresponding to Coded Data from Aurora’s Session 2, item #44 .................................................................................................................. 52
Table 6 Grammar Items Noticed by Type and Session .................................................................................................................. 58
Table 7 Point of Noticing by Category ........................................................................................................................................ 62
Table 8 Grammar, Pronunciation, and Content items at First Noticing ................................................................................. 65
Table 9 Percentage of Trackable items at First Noticing, Product, and Last Word ........................................................................ 66
Table 10 Percentage of Items by Category Linking First Noticing with Product ................................................................. 67
Table 11 Results of chi-square tests locating significant difference by Session and Category for Individuals and Pairs by Point of Noticing ........................................................................................................................................ 76
Table 12 Percentage of Trackable items at First Noticing by Session ................................................................................. 78
Table 13 Percentage of Trackable items for each Category by Session ................................................................................. 80
Table 14 Trackable Correct, Incorrect, and Not Quite Correct Product by Session ........................................................................ 81
Table 15 FN(Correct) + Product (Correct) over Sessions ........................................................................................................ 82
Table 16 Summary of salient findings for RQ 3 ........................................................................................................................................ 83
Table 17 Verbalized and Non-verbalized Items in Grammar, Pronunciation, and Content ........................................................................ 84
Table 18 Non-verbalized and Verbalized Grammar Items and Outcomes .................................................................................. 86
Table 19 Grammar, Pronunciation, and Content items by Engagement ..................................................................................... 87
Table 20 Results of Chi-square Tests Locating Significant Difference by Grammar, Pronunciation, and Content for Individuals and Pairs for Engagement ........................................................................................................................................ 88
Table 21 Grammar: Verbalized and Trackable Data by Engagement for Session 3 .................................................................................. 90
Table 22 Pronunciation: Accept, Deny, and Pronunciation Items Identified by Individuals and Pairs Over the Sessions ........................................................................................................................................ 91
Table 23 Pronunciation: Verbalized and Trackable Items by Engagement for Individuals and Pairs ........................................................................................................................................ 92
Table 24 Verbalized and Trackable Content Items Identified by Individuals and Pairs by Engagement Over Session ........................................................................................................................................ 93
Table 25 Engagement and FN(Correct) items for Grammar, Pronunciation, and Content ........................................................................ 94
Table 26 Grammar: Verbalized and Trackable Data Only by Engagement and Product ........................................................................ 95
Table 27 Engagement and FN(Correct) + Product (Correct) Items for Pronunciation and Content ........................................................................................................................................ 96
Table 28 Engagement and FN(Noticed, but no Clear Decision)+ Product (Correct) Items for Pronunciation and Content ........................................................................................................................................ 96
Table 29 Summary of salient findings for RQ 4 ........................................................................................................................................ 101
Table 30 Individual and Pair Session Recording Time and Length of Transcripts .................................................................................. 106
Table 31 Number of Items Noticed by Extended Example Participants .......................................................................................... 106
Table 32 A Comparison of What was Said During Preparation Time Before PT 2 and PT 3 ................................................................. 114
LIST OF FIGURES

Figure 1. Consciousness in a multi-store model of memory. (Schmidt, 1990, p. 135, after Kihlstrom, 1984). ................................................................. 9
Figure 2. A typical seating arrangement for the Paper D examination. ........................................ 36
Figure 3. Organization of participants into units and assigned condition................................. 41
Figure 4. The stages of a session. .......................................................................................... 43
Figure 5. Examples of items that were marked in a transcript produced by a participant......... 46
Figure 6. Correspondence between coded data and trackable item labels First Noticing, Product, and Last Word.......................................................... 52
Figure 7. Distribution of coded items by category. ................................................................. 57
Figure 8. Type of Pronunciation items identified by proportion. ........................................... 59
Figure 9. Type of Content items identified by proportion. ..................................................... 60
Figure 10. Sorting Judgement Possible items into Trackable and Untrackable by category....... 63
Figure 11. Point of Noticing by percentage for Individuals and Pairs..................................... 77
Figure 12. Differences in Engagement in Content at Session 3 between Individuals and Pairs. . 89
## List of Appendices

Appendix A Sample Examination From the Use of English Paper D ........................................ 160  
Appendix B Pilot Study .............................................................................................................. 164  
Appendix C Questionnaire for Participants at the Start of the Study ..................................... 167  
Appendix D Participant Questionnaire at the Completion of Study ......................................... 168  
Appendix E Summary of Data for Questionnaire Administered at the Beginning of the Study 169  
Appendix F Summary of Data for Questionnaire Administered at the Beginning of the Study 170  
Appendix G Summary of Data for Questionnaire Administered at the End of the Study ....... 171  
Appendix H Code Book ........................................................................................................... 172  
Appendix I Coding Rules ....................................................................................................... 174  
Appendix J Coding for Engagement When Verbalizing ............................................................. 177  
Appendix K Difficulties Encountered During Data Collection .................................................. 179  
Appendix L An Example of an Extended Discussion ................................................................. 180  
Appendix M Glossary .............................................................................................................. 184
CHAPTER 1: INTRODUCTION

My research has its roots in my experience as an English language teacher in an English language Hong Kong secondary school. Most people in Hong Kong speak Cantonese as their first language, and all the ethnic Chinese students I taught were Cantonese speakers. Even the few non-Chinese students in the school were remarkably competent in Cantonese. In this environment, my school stood as an English language oasis. The aim of the school was to immerse the students in an English environment as far as possible. All subjects offered by the school were taught in English with English textbooks, with the exception of Chinese and French language studies. The students had contact with English every day, and were also provided with opportunities to use the language in a range of activities organized throughout the year by an active English panel. The school firmly adhered to English for all official functions, communications, and announcements. Nurtured in this environment, many of my students were comfortable using English to learn, and could speak English well.

There were, however, many who spoke English, and for the most part made themselves understood, with apparent disregard for grammatical accuracy. My curiosity was piqued. Why did this happen? All the senior students had had at least five years in an environment where they were exposed to English every day. With so much experience with English in the classroom and during school activities, I had expected well-formed, grammatical utterances to be the norm. What accounted for some students being more fluent than others? Could it be that many of the students could not “hear” the mistakes they were making? If they were given opportunities to notice their mistakes and repeat their performances, would their performances improve? Would

---

1 The English department was staffed by a mix of native and near-native speakers of English, while members of other academic departments all spoke reasonably good English.
they be better able to notice their inaccuracies working on their own, or with the help of a classmate?

I began looking for answers to these questions in the second language acquisition (SLA) literature. Ortega (2009) explained that according to the cognitive-interactionist tradition, attention is a fifth key ingredient in language learning after input, interaction, language aptitude, and output. Schmidt’s (1983) well-known case study of Wes is an illustration of a second language (L2) English learner who had the right ingredients for language learning: a positive attitude towards the target language, and an optimal linguistic environment, that is, one that provides input, interaction, and opportunities for output. In the short time that he had lived in Hawai’i, having moved from Japan with minimal ability in English, Wes was already using the L2, English, most of the time and in sophisticated ways. His personal charm, creativity, and willingness to communicate enabled him to hold conversations without difficulty. However, after three years, his utterances were still ungrammatical, and he admitted that he spoke “funny English.” What was missing appeared to be attention to grammatical form.

Schmidt and Frota (1986) conducted what was to become another well-known case study, this one in Rio de Janeiro. In this study, Schmidt documented his own learning of Portuguese. He took a 5-week intensive course in Portuguese using the audiolingual approach, but his main language learning was through interaction with native speakers. Based on a journal Schmidt kept of his experiences learning Brazilian Portuguese, and a series of tape-recorded conversations between him and Frota, the researchers made several discoveries. First, they discovered that the forms Schmidt produced in conversation, other than those that had been taught to him explicitly, could be traced to what he had noted in his journal previously. That is, noticing forms led to the production of the forms. Secondly, the recordings showed that he was frequently corrected while in conversation with native speakers. However, this had no effect because he was unaware that
he was being corrected. Thus, the researchers concluded that unless learners consciously compare their own output with correct speech, they will not be able to overcome errors.

While I can draw some parallels between Wes, Schmidt, and my students, there were notable differences. For one, my students did not inhabit an optimal linguistic environment for English language learning. Cantonese was the preferred language of communication in almost all aspects of their lives outside the classroom. Second, my students had been learning English formally for most of their school career. Yet, it appeared that they had come to a point of stagnation in their grammatical competence even as they continued receiving English lessons, a fact that led me to suspect that they had stopped noticing their errors, especially when they spoke. Thus, Schmidt’s conceptualization of and explanation of “noticing” and “attention” is a starting point for my research.

A great deal of research has drawn on Schmidt’s (1990) seminal paper and subsequent elucidations of noticing and attention (1995, 2001). Many researchers have since investigated noticing as a phenomenon, and as a crucial part of language learning. I summarize some of the pertinent research in this area in Chapter 2.

The social turn in SLA (Block, 2003) led me to look for answers to my question in a different direction. On the one hand, as mentioned earlier, the cognitive-interactionist tradition understands SLA as a cognitive process that individuals undertake. In that view, relationships between individual and community, and social and cultural issues are not examined. On the other hand, a Vygotskian sociocultural theory (SCT) of mind focuses on “the relationships between the individual’s physiological aspects and the social and culturally produced contexts and artefacts that transform the individual’s cognitive or mental functions” (Swain, Kinnear & Steinman, 2011, p. xiii).
Recent scholarship using an SCT framework has enriched our understanding of language learning in a number of important ways and given us new ways to think about this phenomenon. For example, language can be conceptualized as a tool to mediate cognition, that is, languaging (Swain, 2006), and the internalization of language can be understood as arising from social interaction. Or one might use a microgenetic method in order to understand language development. These ideas intrigued me and led me to wonder how languaging and collaborative dialogue facilitate noticing. In Chapter 2, I review some of the relevant collaborative dialogue studies in SLA that inspired my study.

I also investigated research into the noticing of language forms and speaking, and found six papers of particular relevance (Lynch, 2001, 2007; Mennim, 2003, 2007, 2012; Stillwell et al., 2010). All of these studies were classroom based, and the investigators used transcription to promote noticing. In producing a written script, the student’s speech was “made visible” for reflection. When learners transcribed their own speech, they noticed the gaps in their own linguistic knowledge. In all cases there were no pre-determined linguistic features selected by the researchers as a target of the study. The research involved learners working with a peer or peers, and opportunities for noticing were created through transcribing and editing with a fellow learner. As well, students were able to make use of additional feedback from the teacher, or hold discussions with the teacher, which represented another opportunity to notice and engage in languaging (Swain, 2006). In all six studies, student feedback on the transcription activity was found to be positive, and the effects of the noticing activity were durable. The researchers acknowledged that a repetition effect could have accounted for the subsequent improvement in production. Notably, the activities arose as an extension of what was already in the curriculum. I review these studies in greater detail in Chapter 2.
Rationale

I am pursuing this research for a number of reasons. First, I wish to investigate the importance of noticing and attention to language forms described by Schmidt’s noticing hypothesis (Schmidt, 2001). Many researchers (Gass, 1997; Schmidt, 1990; Skehan, 1998; VanPatten, 1990) agree that noticing is a key element in L2 acquisition. In addition considerable research has been done to investigate the ways in which attention to form can be encouraged and how this may lead to uptake (e.g., Alanen, 1995; Harley, 1998; Leeman, 2003; Lyster & Ranta, 1997; Mackey & Philp, 1998; Oliver, 1995, 2000; Philp, 2003; Swain & Lapkin, 1998, 2002; White, 1998). Some of the ways noticing has been encouraged and investigated are implicit or explicit instruction (DeKeyser, 2003; Ellis, R., 2009), consciousness raising through input enhancement (Leow, Egi, Nuevo, & Tsai, 2003; Sharwood Smith, 1993; White, 1998), and explicit or implicit rule giving (Robinson, 1997). I wondered if using transcription of their oral production as a route to noticing could help the students I taught to speak more grammatical English. My study is intended to add to the existing literature on noticing and speaking.

Instead of focusing on one particular language feature, I wanted to know what it is that learners pay attention to, and what outcomes result from this noticing if they have a chance to revisit their speech production. When speaking, there is little time to plan or edit. Using retrospective procedures, my goal was to collect evidence about what learners notice about their language production, and what happens to the items that have been noticed.

Also, this research adds to our understanding of secondary school English learners in Hong Kong. About 60,000 students prepare for university entrance exams annually in Hong Kong, and since English is a compulsory subject, all students must take a high-stakes English examination. This study is designed to investigate how students identify and self-repair their mistakes alone or in collaboration with a peer partner, and determine whether Hong Kong
students are able to make use of learning opportunities that are not teacher-centred. It also illustrates that simple technologies, such as audio-recordings, can be used as pedagogical tools for language learning in general and spoken language learning in particular.

To elicit oral presentations, I used past-year Hong Kong university entrance examinations, the format of which was familiar to the students. Participants worked individually or in pairs, and I built transcription, talk-aloud and peer interaction opportunities into the research design. I also collected background language profiles, and conducted an interview with each participant at the end of the data collection period. I elaborate on the research design in Chapter 3.

Organization of thesis

Chapter 2 presents a review of the literature relevant to the research questions under investigation. The literature informs the research design and the data analysis that follows. This study is motivated by several questions. First, I wanted to know what students would notice if they were given the opportunity to closely examine language they had produced. Would transcribing their own spoken language help them to notice their own errors? If they were given a chance to do the same oral task again, would they modify it based on what they had noticed? And if this exercise of noticing was repeated over the course of a semester, could improvements in their mastery of the language, if any, be traced? Would how and what they noticed be affected by whether they were working alone or in collaboration with a peer? These questions are distilled and presented as the research questions for the study in Chapter 2.

Following this, in Chapter 3, the context in which the study was conducted is described. Also there how the research was conducted, the procedures used, and how the data were collected and processed is described. In Chapters 4 and 5, I present and discuss the results of the
quantitative analysis of the data with respect to the research questions. In Chapter 6, in order to provide a deeper understanding of the phenomenon under study, I present data qualitatively. In Chapter 7, I discuss the study in light of previous studies, and look at the pedagogical applications and limitations of the study. I conclude with suggestions for further investigation.
CHAPTER 2: LITERATURE REVIEW

This review of the literature is divided three sections. In the first is a review of the literature on the Noticing Hypothesis (Schmidt, 1990, 1995, 2001) and some of the research studies associated with it. In the second is a review of a selection of the SLA research into collaborative dialogue that is relevant to this study, while in the third is a review of studies that have looked at the use of transcription as a way of helping learners to notice language forms. Together, this literature informs my research, the design of the study, and the analysis I carried out.

Section One: Noticing

In the SLA literature, ‘noticing’ is attributed to Schmidt’s (1990) seminal paper, “The role of consciousness in second language learning.” Based on theories of information processing in cognitive psychology, Schmidt (1990) stated that “almost all theories of consciousness attempt to explain what I have called noticing and the systems that give rise to it” (p. 135). In general, information processing theories assume that human memory, attention, and processing capabilities are finite. Multi-store theories describe memory as consisting of three components in general: sensory registers, short-term memory store, and long-term store. This is illustrated in Figure 1.
In order for learning to take place, new information must be “noticed” or made the subject of focal attention. According to this model, information is lost if it is not attended to. If the information is noticed, it will be encoded in short-term memory for further processing or rehearsal. If no further processing takes place, this information may be forgotten. Information that has been processed will be “learned” and subsequently stored in long-term memory so that retrieval is possible. In this model of memory, the role of attention is explicated through the use of various metaphors such as a “switchboard,” “gate” or “filter,” in other words as a device that protects us from sensory overload. (Broadbent, 1982; Posner & Klein, 1973; Posner & Snyder, 1975; Kahneman & Treisman, 1984; all cited in Schmidt, 1990, p.136).

Attention thus controls what is processed and subsequently learned. Building on this orthodox cognitive psychology model of attention and learning in input processing, Schmidt (1990) defined “intake” as “that part of the input that the learner notices” (p. 139). He went on to state that “if noticed, it becomes intake.” Thus, he put forth what is now considered the strong version of the Noticing Hypothesis, “noticing is the necessary and sufficient condition for converting input to intake” (p. 129).
As the hypothesis gained currency, Schmidt (2001) refined and clarified his statement on attention and awareness. In his discussion he limited noticing to a low level of abstraction, stating that:

‘Noticing’ is therefore used … in a restricted sense, as a technical term equivalent to ‘apperception’ (Gass, 1988), to Tomlin and Villa’s (1994) ‘detection within selective attention,’ and to Robinson’s (1995) ‘detection plus rehearsal in short-term memory. My intention is to separate ‘noticing’ from ‘metalinguistic awareness’ as clearly as possible, by assuming that the objects of attention and noticing are elements of the surface structure of utterances in the input – instances of language, rather than any abstract rules or principles of which such instances may be exemplars. (Schmidt, 2001, p.5)

Instead of “noticing,” he advanced “attention” as the construct that “appears necessary for understanding nearly every aspect of second and foreign language learning” (p. 6).

**Attention in SLA research**

Schmidt (2001) noted that virtually every aspect of SLA and foreign language learning depends on some understanding of attention (p. 6). Attention is what allows learners to notice the “hole” (Doughty and Williams, 1998) or the gap between what they can produce and what they need to produce. Lack of attention has been pointed to as the cause of the variations between fossilized end-state L2 learners and more proficient L2 learners (Han & Selinker, 2005).

Attention plays a role in the development of fluency because it facilitates successful encoding of information in memory. When this information is used frequently, it becomes automatic to the learner (DeKeyser, 2003). It is also an important construct in explaining differences between learners. Gardner and Tremblay (1995) proposed that attention is one of the motivational behaviours that accounts for higher proficiency or better performance. Strong links have been found between motivation and learning strategies, particularly cognitive and metacognitive strategies, which include strategies for focusing or sustaining attention (Schmidt et al., 1997). In addition, attention can explain differences in the rate new information is encoded in short-term
memory, and why some learners have greater language aptitude in terms of their ability to take in, recall, and analyse the input they receive (Schmidt, 2001, p. 10).

Most importantly, however, attention has been the construct that has driven a great deal of research in language instruction, both in classroom (Doughty & Varela, 1998; Harley, 1998; Lyster & Ranta, 1997) and laboratory settings (de Graff, 1997; Robinson, 1997), particularly research into the ways in which learner attention can be drawn to a language feature in order to produce more accurate output. Many of these studies take as a point of departure Long’s (1996) revised Interaction Hypothesis, in which he stated that selective attention is a mediator that clarifies the process whereby input (from “environmental contributions”) becomes intake, and allows learners to notice. Many studies have been done, ranging from input enhancement or input flood (Alanen, 1995; Izumi, 2002; White, 1998); to input processing (VanPatten, 1990; VanPatten & Cadierno, 1993; Wong, 2001); to interactional recasts (Ellis, Basturkmen, & Loewen, 2001; Iwashita, 2003; Leeman, 2003; Lyster, 1998a, 1998b; Mackey & Philp, 1998; Oliver & Mackey, 2003; Philp, 2003); to studies on output (Adams, 2003; Izumi, Bigelow, Fujiwara & Farnow, 1999; Izumi, 2002; Swain & Lapkin, 1998, 2002). In their reviews of the research on form-focused instruction, Norris & Ortega (2000) and Spada (1997, 2011) have noted that drawing attention to form, notably through explicit instruction and corrective feedback, facilitates language learning. Pedagogical intervention is required (Doughty, 2003) because some of aspects of the L2 are not noticed because they are, for example, conceptually dissimilar to the L1, not communicatively problematic, or not acoustically salient (p. 291). In corrective feedback research, noticing offers an explanation of why recasts are sometimes not perceived as corrective feedback by learners (Nicholas et al., 2001). Further, what is noticed may not be perceived in the way intended by researchers, and is consequently not attended to (Mackey, Gass, & McDonough, 2000). Thus, attention and noticing are key constructs in these
studies. But little of how attention actually works is explained in these studies, prompting Schmidt (2001) to note that attention is merely a *deus ex machina* in these accounts (p.11).

As most of the research on noticing, awareness, and attention in SLA has been conducted within an input-processing model of learning, which draws on cognitive psychology, the investigations have also mostly followed procedures from an experimental or quasi-experimental paradigm, which uses pre-test, post-test and delayed post-test treatments to evaluate the efficacy of a treatment in facilitating SLA. These studies are researcher-driven, the learners who participate are largely seen as processing units. The researcher determines a specific language feature to be the focus of treatment, and measures learner performance on the assumption that the outcome is a result of the treatment. What Schmidt (1990) noted early on is telling:

> The primary reason that we have undervalued the role of consciousness in second language learning is that we have simply not done much research to assess sensitively what learners notice and what they think as they learn second languages. Like behaviorists who assumed that their subjects left their mental faculties outside the laboratory door, we have assumed learner ignorance more often than we have attempted to investigate learner awareness. (Schmidt, 1990, p. 150).

In one of a series of studies on Spanish learners conducted to investigate whether learners were aware or unaware of the irregular third person singular and the plural preterit forms, Leow (2001) sought to understand how the participants were processing input by using think-aloud protocols. The stimulus was a crossword puzzle designed to focus attention on particular forms. As learners did the task, they were asked to “think aloud.” These exercises were tape-recorded, transcribed, and coded as to whether the learner was “aware” or “unaware.” The researcher noted that despite controlling the variables in his study, and having all the participants perform the same task, learners demonstrated differences in processing, indicating that “in addition to type of task driving a specific type of processing, learners’ individual choice might also do so” (p. 141). But what kind of choices were the learners making, and what was causing them to make such
choices? Leow also noted that though the use of online data has higher internal validity, it drew attention to the fact that learners were acting differently, citing his own study (Leow, 2000) in which half of his participants exposed to the same task with the same instructions performed differently from the other half in the way they processed L2 data. He considered this a “disturbing finding,” and it led him to wonder how representative his participants were in the different experimental groups (p. 115). While the information processing model of learning is useful for accounting for learning in SLA by examining psycholinguistic factors, there remain areas which are unaccounted for in this paradigm.

In *The Social Turn in Second Language Acquisition*, Block (2003) critiqued cognitively oriented studies, such as Mackey, Gass and McDonough’s (2000) study, but his concerns were not restricted to their study. One of the goals of the Mackey et al. (2000) study was to discover if learners perceived interactional input accurately. Ten ESL and seven Italian as a foreign language (IFL) learners worked in dyads with a native or near native speaker of the target language, who interacted and gave them feedback. Each dyad was videotaped. Immediately afterwards, learners were asked to recall their thoughts as the videotape was played back to them. The researchers found that learners were generally accurate in their perception of feedback about lexis and phonology, but they were generally inaccurate in their perceptions regarding morphosyntactic feedback. ESL learners perceived morphosyntactic feedback as being about semantic content (about meaning in general), but not about morphosyntax at all. IFL learners perceived morphosyntactic feedback as being largely about lexis, operationalized as specific comments about a known or unknown word. The researchers suggest that the reason why morphosyntactic feedback was not perceived as such could be because these aspects of language were not crucial for understanding the interaction. Stimulated recall protocols were used, but these were only coded for accurate or inaccurate perception of the interactional input.
In his critique of this study, Block (2003) suggested that the data could be read and mined more richly to give us a fuller picture of what learners were actually perceiving, or attending to, and why they were producing what they produced at a given time or place. One example he raised was a learner saying in the stimulated recall that “I was thinking that my pronounce, pronunciation is very horrible” (p. 86). Mackey et al. (2000) noted this as an example of the learner accurately perceiving the feedback as phonological. But Block (2003) contended that the learner was also saying much more. In trying to save face, the learner made this statement to tell the researcher that he was aware of his poor pronunciation, and in dealing with his embarrassment, he made the self-effacing remark. To Block, the episode was “an occasion for the participant to negotiate different aspects of his identity (deficient communicator and self-aware) as well as to save face” (p. 86-87). The study is illustrative of the line of inquiry carried out within what Block (2003) terms the “Input Interaction Output” (IIO) model, or one which follows from a research tradition that operates on the premise of Cartesian mind-body dualism. The learner is essentially seen as a processor of information.

Historically, the cognitivist approach has dominated the field of SLA; however, there is growing recognition that other approaches in SLA can not only enrich our understanding of language learning, but also provide a fundamentally different orientation for conceiving how language is acquired (Atkinson, 2011). One influential view of language acquisition is through the lens of sociocultural theory (SCT), which is founded on the work of L. S. Vygotsky (Lantolf & Thorne, 2006). Vygotsky did not conceive of human minds as stand-alone processors that became incrementally more adept and powerful as they processed external stimuli. Instead, he proposed that in humans, biological and cultural factors form a dialectically organized mental system in which biology provides the necessary functions and culture empowers humans to intentionally regulate these functions “from the outside” (from Vygotsky, 1997, p. 55, in Lantolf,
Such a view places a great deal of importance on the social, cultural, and historical context of the learner, but it also resolves what is termed “downward reductionism,” which describes mental processes strictly as a function of biology, with “upward reductionism,” which sees mental functioning as a result of strictly social-discursive factors, by presenting a new ontological view of humans as mediated beings (Lantolf, 2006, p. 69). The way that culture empowers humans “from the outside” is through mediation, a fundamental concept in SCT (Lantolf, 2000, p.1). Thus, for internalization to occur, the activity must be mediated, and, as Swain et al. (2002) stated, language is one of the most important tools for mediation (p. 172).

SCT perspectives complement our understanding of SLA gained from input-processing research paradigms. As Lantolf and Pavlenko (2001) aptly noted,

> Learners have to be seen as more than processing devices that convert linguistic input into well-formed (or not so well-formed) outputs. They need to be understood as people, which in turn means we need to appreciate their human agency. As agents, learners actively engage in constructing the terms and conditions of their own learning. (p. 145)

To date, many studies have been conducted using an SCT approach in investigating language learning, and published collections of these papers have appeared in Lantolf & Appel (1994) and Lantolf (2000), and textbooks on SCT in L2 education (Swain, Kinnear & Steinman, 2010), teaching (Lantolf & Poehner, 2011) and the genesis of second language (Lantolf & Thorne, 2006) have also been published. The areas under investigation are varied and shed light on many areas, including: private speech (Borer, 2004; Ohta, 2000); internalization and regulation (Aljaafreh & Lantolf, 1994); negotiated feedback (Nassaji & Swain, 2002; Nassaji 2011); L1 use in L2 learning (Anton & DiCamilla, 1998; Swain & Lapkin, 2000); collaborative dialogue (Donato, 1994, 2000; Swain and Lapkin, 1998, 2001; Swain, 2000), among others.

Thus, from a sociocultural theory perspective, collaborative dialogue studies are a rich source of data for understanding mediation and noticing. However, collaborative dialogue studies...
have also been carried out from an interactionist-cognitivist perspective. Below, a description and discussion of both types of research are provided.

Section Two: Collaborative Dialogue

Swain’s “output hypothesis” (1993) proposes that “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” (p. 373, Swain (1995)). Examples of research that focused on learner output informed by an information processing perspective include Adams, 2003; Izumi, Bigelow, Fujiwara, and Fearnnow, 1999; and Izumi, 2002. In these studies learner output performs a “noticing” or “triggering” role. The output hypothesis was extended in Swain (2000) beyond the initial frame of reference within an information processing perspective to a consideration of collaborative dialogue and its role in mediating language learning. 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 speak, 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. (2000, p. 113).

As language itself is the cognitive tool that mediates learners’ understanding and holds their attention, the use of language in collaborative dialogues enables us to observe how collaborative dyads work through the problem at hand, what they attend to and why, and how they arrive at the solutions, be they correct or incorrect. Verbalizing makes visible the problem that is being solved, and provides the platform for further reflection in the form of externalized dialogue (Swain & Lapkin, 2001; Tocalli-Beller & Swain, 2005; Borer, 2005; Holunga, 1994; Storch, 2002). Donato (1994) found that collaborative pairs internalized the scaffolded help they
received from each other. Swain and Lapkin (1998) demonstrated that the site of learning was in the interaction between the pairs as they produced language (output) and sought to resolve their difficulties while performing their task. Thus, as stated in Swain and Watanabe (2013), “the importance of collaborative dialogue to teachers and learners of second or foreign languages is that by listening to it, insights into the cognitive and affective processes that learners are using in their particular learning teaching context are inevitable.”

Collaborative dialogue research studies span the four skills: speaking (e.g., Lynch 2001, 2007), listening (e.g., Garcia and Ascención, 2001), reading (e.g., Cross, 2010), and writing (e.g., Storch 2002), with the greatest number focused on the latter. In the L2 context, Storch (2013) noted that most studies comparing collaborative and solitary writing sought to discover whether the collaborative condition produced more accurate performance (p. 72). Two studies reporting greater accuracy for learners working collaboratively over learners working individually are Nassaji and Tian (2010) and Reinders (2009) (in Storch, 2013, p. 73). Nassaji and Tian (2010) studied low-intermediate ESL learners who completed cloze and editing tasks that targeted phrasal verbs, while Reinders (2009) conducted a study with high-intermediate ESL learners collaborating on a dictogloss task that targeted negative adverbs. One reason that pairs performed better could be the amount of time they spent on the task. Storch (2007) found that pairs took longer to complete their tasks, and surmised that collaborative pairs exert pressure on each other to complete a more accurate text. However, it is not simply the amount of time that the pairs take that improves their performance. The quality of engagement in the language-related episodes (LREs) is also an important consideration in discussing the gains made by learners in collaborative dialogue studies. As expected, where there was extensive and elaborate engagement, it tended to be accompanied by LRE resolution (Storch, 2007). Kim’s (2008) study connected LRE resolutions of the targeted vocabulary items to performance on post-tests, and
found that pairs outperformed individuals because they resolved the LREs correctly in most cases, in contrast to individuals, who resolved incorrectly or not at all over half of the LREs.

There are, however, studies reporting no benefit for collaborative dialogue. Kuiken and Vedder (2002) found no benefit for collaborative writing in language learning in a comparison of group and individual task completion. In this study of Dutch high school students performing the dictogloss task, the learners tended to avoid using the targeted passive verb structures. In the analysis of the interaction data for LREs, the researchers found that there was lengthy and elaborate noticing in some cases, as well as mere mentions of the passive structures, but the researchers did not link the amount of noticing to the results of the post-test.

Collaborative dialogue studies have also investigated whether the partnership of the dyad affects learning outcomes. Watanabe and Swain (2007) found that the highest scores on the task were obtained by participants in a collaborative relationship, or where the participant assumed the role of the expert in an expert/novice relationship. These collaboratively oriented pairs identified more LREs than other pairs, and generated more discussion of language items. Conversely, dyads in a dominant/novice relationship were not as successful at learning. However, Storch (2013) contended that a more important factor in the success of a collaboration is how learners orient to the activity (p.69), that is, what learners’ attitudes and goals are.

Storch (2013) reported that task type and the L2 proficiency of the learners could have an effect on the number of LREs that are generated, their type (lexis or form), the quality of the learners’ engagement, and the outcomes (p. 70). Since the generation of LREs is a measure of attention to form, for the same type of task, studies have shown that higher proficiency learners generate and engage in a higher number of LREs (Leeser, 2004; Swain & Lapkin, 1995). As well, some studies have raised the question as to whether there are benefits for pair work among low-proficiency learners and whether they have the linguistic resources to resolve the LREs they
identify (Leeser, 2004). Swain and Lapkin (1995) noted that the young participants in their study were unable to verbalize the difficulties they were facing, but could only say that something was non-native-like, relying largely on “sound” and “sense.” The researchers suggested that that could be a result of the participants’ experience of learning French in highly communicative and experiential immersion.

To find out if the gains made in collaborative activity endure, a longitudinal study was conducted by Shehadeh (2011) in the United Arab Emirates. The participants were low-intermediate EFL university students who were divided into two groups. One group wrote individually, while the other group wrote in pairs. The pairs were changed every two or three weeks. The learners completed 12 pieces of writing over a period of 16 weeks. When the groups were compared, Shehadeh found that both groups of students had improved in their writing, but paired students showed greater improvement in content, organization, and vocabulary than individual students. However, there was no improvement in the grammar and mechanics of the writing of the students who collaborated. Though learners were pooling their knowledge, deliberating about grammar was challenging. It seemed that the low-proficiency learners were unable to mediate the gaps in their grammar knowledge without input from more knowledgeable sources. Overall, Shehadeh discovered that learners in pairs viewed collaborative writing positively, and gained confidence as writers as a result of the activity. This result is consonant with Yarrow and Topping’s (2001) longitudinal study of young children writing in pairs or individually in their L1. Yarrow and Topping found that paired writers were better able to keep each other on task. As well, collaboration provided opportunities to explore ideas and to seek or provide help to one another. Through collaboration, the pairs gained confidence in writing, and they reported enjoying co-authoring.
While transcription was employed in processing the data in most of these collaborative dialogue studies, it was not performed by the participants. In the next section I review studies that involved participants’ self-transcription.

Section Three: Transcription Studies

One way in which noticing has been examined in the research literature is through the use of self-transcriptions. In Lynch (2001), this was used as an opportunity for learners to reflect on their own linguistic output. This study was done in the United Kingdom. The eight learners were from of a university English for Academic Purposes (EAP) oral communication skills class, they came from six different countries, and their average proficiency was approximately 5.5 IELTS (p. 126). Typically, these students would take part in classroom activities that included role-plays performed in pairs in front of the class and videotaped. As an extension to the classroom activity, the students who were invited to participate in this study were asked to transcribe a self-selected extract of between 90 and 120 seconds of a recording of a role-play they had performed in front of the class. Each student involved produced a transcript of the extract, and the pair had to agree on a final version, which was called Transcript 1. After that, the researcher asked them to make changes to Transcript 1 until they were satisfied with the language. As they reviewed and made changes, they were videotaped. The product of this process was Transcript 2. Lynch noted that the learners cooperated with each other, and were motivated to do their best. He also noted that “the time and trouble they take over details is striking (p.128).” For each pair that he videotaped, the 45-minute tape ran out before they had completed the transcription and the activity. In the next stage, the students submitted both Transcript 1 and 2 to the teacher who corrected Transcript 2, changing only the parts that were linguistically inaccurate or expressed in a manner that was
marked. This revised piece, or Transcript 3, was returned to the learners as feedback, and the teacher and learners discussed various points further (p. 126 - 127).

The study found that the students were motivated and cooperated well with each other as they worked on Transcript 1 and Transcript 2. The four pairs noticed on average 28 points for change in the transcript, which is about one change for every four seconds of speech (p. 128). Lynch used five categories to catalogue the changes: grammatical correction; lexical correction; editing; reformulation; and mixed. Most of the changes were for the better. Lynch found that changes were initiated almost equally between the original speaker and the original listener. He observed that three of the pairs seemed to share responsibility for initiating changes, while the fourth pair behaved differently, one member of that pair assumed a dominant role. He suggested that this could have been a result of the gender difference, as this was also the only mixed pair, or of cultural differences between the Arab and Mandarin speaker. Both the students and the teacher made changes in grammar, particularly in verb tenses and article choice. While the students also made a substantial number of changes categorized as editing and reformulation for greater precision, they made relatively few changes to vocabulary compared to the teacher (p. 129).

Though the students had been given time to notice their errors and make changes to their transcripts, there were still areas that needed to be corrected. In all, Lynch noted that students noticed and modified about 60% of the points in need of correction, while he changed the remaining 40%. Lynch concluded that the transcribing activity yielded several opportunities for noticing: the transcribing process, reflective self-correction, collaboration with a peer, and finally, teacher feedback. He considered this combination to be the optimal mix of feedback in classroom teaching (p.130).

Lynch (2007) conducted a follow-up study in the same university EAP setting in the UK. The stated aims of this study were to explore “the feasibility of managing self-transcribing and
editing in class, and the potential learning benefits of noticing tasks based on transcripts” (p. 311) under ordinary classroom conditions. Once again, the material for the study was generated from the role-plays that were performed and videotaped by the students in the classroom. However, there were several differences between this study and the one previously described. The main difference was in the sources of the changes to the pairs’ output: either student-initiated (SI), or teacher-initiated (TI). Another difference was that twice the number of students were involved in this study (n=16), eight from one class and eight from another. The eight students from Class 1 ranged in proficiency between 4.5 and 6.5 in IELTS Listening, while the eight from Class 2 ranged between 6.0 and 7.5 in IELTS Listening. The students came from 10 different countries.

The study was quasi-experimental in design. Class 1 students followed the SI procedure, where the students selected a portion of their own role-play performance to transcribe. Class 2 students received a series of extracts transcribed by the teacher, who had selected the portions of the students’ role-play that he felt were problematic. The pairs then corrected the transcript before checking their changes with the teacher. In order to investigate whether there were long-term gains, students were asked to perform and record the role-play again on the day that they were given back the teacher reformulated transcript, and again four weeks later. These recordings were transcribed by the researcher and compared with the earlier performances.

Lynch found that “self-transcribing could be managed as a routine activity in the classroom” (p. 314), and that in the TI procedure, the work of transcription and correction was not excessive for the teacher. Students who had followed the SI procedure achieved a higher percentage of accuracy after a four-week interval (64%) than the students who had followed the TI procedure (47%). Lynch argued that the difference was a result of differences in depth of processing. Those following the SI procedure had greater opportunities for self- and peer-
correction as they worked in collaboration with another to produce Transcript 1, while students following the TI procedure did not have the benefit of discussion with or correction by a peer.

Mennim’s (2003) study was conducted in Japan with three first year university students. These students were in an upper-level English course and had TOEFL scores ranging between 500 and 550. In their course, the students were required to give an in-depth presentation on a topic of their choice. The aim of the study was to determine whether students could take advantage of a rehearsal in order to make improvements to their spoken output in the final presentation (p. 133). The three students tape-recorded a 20-minute rehearsal of their presentation with the teacher as the only audience. No scripts were used, but the students were allowed to use small cue cards. After that, the students selected and transcribed verbatim a 5-minute segment that contained equal contribution from each of the group members. This transcript was typed, and the students corrected their own parts in the transcript with a red pen. This was handed to the teacher who added corrections for errors they had missed. The transcript was returned to the students one week before the final presentation. The following week, the students did their final presentation, which was also tape-recorded. The teacher transcribed the section of the final presentation that corresponded to the part that the students had transcribed from the rehearsal, and compared the two transcripts.

The students noticed 49 possible errors in the transcript of 5 minutes. Since they contributed about equally, each student averaged 16 errors. This was fewer than in Lynch’s (2001) study. In comparing the transcripts of the rehearsals and final presentations, Mennim organized the changes into the following categories: pronunciation, grammatical forms (articles, prepositions, and passive forms), and changes to content. He found that one student made a large number of target-like repairs to articles, but not the other two students. For prepositions, the students seemed to notice the changes and make mostly target-like repairs. For passive forms, the
student who made two changes recalled both in the final presentation, while another one responded to the teacher’s feedback accurately. Mennim noted that the students tended to notice instances of words they had difficulty pronouncing in the rehearsal, and these were produced almost target-like in the final presentation. Where content was concerned, he noted that students made changes to improve the comprehensibility and sophistication of the script. One student made nine changes on the rehearsal transcript, and included all the changes in the final presentation. Another one made two changes, but neither appeared in the final presentation.

From these results, Mennim concluded that the transcription activity had successfully focused student attention on form. He observed that students seemed to have different focuses when doing the transcription. One student focused on linguistic features while transcribing, but seemed to have used the noticing opportunity to improve and expand on the comprehensibility and sophistication of the voice in her presentation. Meanwhile, another student focused exclusively on language. He suggested that perhaps the difference in proficiency (50 points in the TOEFL) between the two students might explain this difference. The one with greater proficiency might have as a result been more able to focus on elements beyond linguistic features.

In a later study, Mennim (2007) focused on two students from the 2003 study and traced the development of language form from the effects of noticing oral output. In the body of noticing data generated by the students as they prepared their transcripts, when an L2 form was noted and discussed, he would he trawl the output data to find subsequent instances of that form. He acknowledged that “one of the inevitable restrictions associated with this procedure was that there was no guarantee that forms noticed would ever re-emerge as output during the presentations” (p. 272). Only seven examples that fulfilled the criterion of appearing over 10 times in the presentation recordings were identified: the non-count noun “garbage” was one of
them. He traced the use of “garbage” by the two students, who were researching and presenting on the topic of garbage disposal, over the course of 6 months. The two participants, Toru and Katsu, noticed the item three and five times over the year respectively. The word was produced by Toru 13 times over the three presentations and one rehearsal, while Katsu used it 25 times over the same period. Mennim found evidence that long-term gain in accuracy could be attributed to the noticing that occurred during the course. Because of the longitudinal nature of the study, it was possible to trace the development of the students’ knowledge, from declarative understanding but only 50% accuracy rate, to 100% accuracy rate 4 months later.

Mennim (2012) showed how his students negotiated language form as they worked on their transcripts. From excerpts of their negotiations as they produced a transcription of their presentation, Mennim illustrated how students used various strategies and resources to complete their task, which included their L1 (Japanese), dictionaries, metalinguistic knowledge, comparisons with existing forms they knew, formulating hypotheses about how L2 forms work, and applying these hypotheses to subsequent points (p. 55). However, Mennim said about how frequently they employed these strategies. He noted that the students were not able to resolve every problem, but this did not result in frustration, because they were aware that the teacher could be consulted at a later time. Thus, the negotiations were helpful because the students were working on forms that were relevant to them, and they served as a catalyst for further noticing of the input. Mennim pointed out that the transcription exercise allowed students to spend more time examining language forms and use in one single lesson than they would have had the teacher been involved in each discussion. Thus, learner negotiation and pooling of knowledge was beneficial in adding to the learners’ L2 knowledge, and the learners took responsibility for correcting L2 errors without the aid of the teacher. Mennim concluded by recommending self-
transcription as “an effective way of generating discussion about language and encouraging learners to think about their own language use” (p. 61).

Stillwell, Curabba, Alexander, Kidd, Kim, Stone, and Wyle (2010) reported on a team of instructors using self-transcription in a poster presentation task in class. The poster presentation task follows from an earlier study done by Lynch and Maclean (2000) which had used this activity. The aims of this task were pedagogic: to promote noticing and reflection on language use during the performance of a task; to provide an engaging alternative to typical end of unit presentations; and to reap benefits in fluency, accuracy, and complexity. The 20 students were Japanese university students in the freshmen English programme who had been placed in the second highest tier based on an in-house placement test. The students made posters summarizing the main points of a challenging text on film genres. The students were divided into pairs, and each student responded to questions about his or her poster posed by the other member of the pair for 3 minutes. The conversations, two for each pair, were recorded, and each student transcribed both conversations. They then worked together to produce corrected versions of both conversations. For homework, the students typed up the transcript of the discussion their own poster only, and submitted it to the teacher for correction. Once the teacher returned the transcripts, the students rejoined their partners to discuss the corrections, and then they had three days to review them. One week later, the entire task was repeated, but with a new partner. By the end of the cycle, the student would have produced two typed transcripts for each poster, containing their own and the teacher’s corrections. A post-task questionnaire in which the students rated the usefulness of each stage of the process was given.

Stillwell et al. (2010) found that transcribing their own oral output provided a valuable means for students to gain insight into their own language development. Students were able to transcribe their work faithfully on the whole, and even more accurately the second time around,
but there were misrepresentations, missing words, and even substantial portions of the recording missing from the transcripts sometimes. Word count as a function of fluency did not yield conclusive information: 45% said more the second time, while 55% said less. The students were aware that in trying to express themselves more clearly, they might use more words to clarify, or say fewer unnecessary things in order to be more direct. One student astutely noted that what was said depended on the quality of the interaction with the interlocutor.

Stillwell et al. used Lynch’s (2001) categories to catalogue the errors. In terms of accuracy, they found that the students improved in their ability to identify errors and correct them. Close to half (47.8%) of the corrections were for grammar, while the rest were reformulations (18.6%) and edits (24.3%). They also tended to notice errors in plurals and subject-verb agreement. When the form that had been corrected was re-used, accuracy was about the same at 55%. With regard to complexity, the researchers found that transcription and repetition led to a 42% increase in complexity in the 53 instances of comparable utterances that were identified. In all, the students had a positive impression of most aspects of the task. They valued their teacher’s corrections most, 92% rated them “very useful” while self-correction was described by 84% of the respondents as “very useful.” They also found transcribing themselves and repeating their presentation useful, but transcribing their peer’s work or correcting their peer were considered less useful. A summary of the studies described in this section is shown in Table 1, on the next page.

In sum, a number of elements were common among these six studies. First, all the studies were pedagogically focused. They were designed with the classroom in mind so that the transcribing activity fit into an existing curriculum. Lynch (2007) answered the question about the feasibility of using transcription in the classroom. His study addressed the amount of time required to include transcription within normal classroom activities, and the effect this had on
teachers’ work load. Stillwell et al. (2010) reported on a transcribing activity integrated within the oral skills part of a curriculum, and concluded with additional suggestions for planning such exercises in the classroom, such as the provision of supplementary lessons and extension activities. Second, none of the studies were focused on a pre-determined linguistic feature as a target. Mennim’s (2003) aim was to increase the focus on language form because communicative tasks were sometimes completed in ungrammatical, but broadly intelligible language.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Role-play.</td>
<td>1. Role-play.</td>
<td>1. Rehearsal of presentation on self-selected topic in groups of 3.</td>
<td>1. Improvised presentation of poster in response to questions about their poster: 3 minutes. Afterwards, the pair switches roles.</td>
<td></td>
</tr>
<tr>
<td>2. Self-selected extract (90-120 sec.) which pair transcribes to produce Transcript 1.</td>
<td>2. Student-initiated (SI) group transcribes self-selected portion of role-play to produce Transcript 1.</td>
<td>2. Transcribe 5 minutes of presentation to produce Transcript 1. Then students mark up the mistakes and correct them.</td>
<td>2. Pair transcribe both 3 minute conversations, and make corrections. This is Transcript 1. Teacher gives feedback.</td>
<td></td>
</tr>
<tr>
<td>3. The pair negotiates changes to Transcript 1 to produce Transcript 2.</td>
<td>3. SI group negotiates changes to Transcript 1 to produce Transcript 2. Teacher-initiated (TI) group receives extracts, and corrects these to produce Transcript 2.</td>
<td>3. Teacher corrects and improves Transcript 1. The product is Transcript 2. This is returned to the group.</td>
<td>3. Pair reviews teacher feedback.</td>
<td></td>
</tr>
<tr>
<td>4. Teacher reformulates Transcript 2, and the product is Transcript 3.</td>
<td>4. Teacher reformulates Transcript 2 =&gt; Transcript 3.</td>
<td>4. One week later, the group does their presentation again.</td>
<td>4. Three days later, repeat entire process with new partner to produce Transcript 2. Teacher gives feedback.</td>
<td></td>
</tr>
<tr>
<td>5. Discussion with teacher comparing Transcript 2 with Transcript 3.</td>
<td>5. Re-record role-play. Compare Transcript 2 with Transcript 3 with teacher.</td>
<td>5. Teacher transcribes corresponding portion of the presentation and compares it with the earlier version.</td>
<td>5. Self-assessment of fluency, accuracy and complexity.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Averaged 28 changes per pair.</td>
<td>• SI group achieved a higher percentage than TI group for accuracy in re-recorded role-play on the points. reproprocessed in Transcript 2: 64% vs. 47% respectively.</td>
<td>• Average 16 errors identified per student.</td>
<td>• Fluency: varied.</td>
<td></td>
</tr>
<tr>
<td>• Mainly grammar 35.7%.</td>
<td>• Students had different foci: articles; elaboration of content; pronunciation.</td>
<td>• Accuracy: uptake of student and teacher corrections. about equal (about 55%).</td>
<td>• Accuracy: uptake of student and teacher corrections. about equal (about 55%).</td>
<td></td>
</tr>
<tr>
<td>• Changes equally identified by speaker and listener.</td>
<td>• Complexity: 3 times more complex, for comparable structures.</td>
<td>• Complexity: 3 times more complex, for comparable structures.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Third, all the learners in the studies were adults. They were enrolled either in a university EAP course, or in a first year university English language course with some focus on oral communication skills. Fourth, all the studies used self-transcription as a means to focus students’ attention on their own oral output. In producing a written script, the student’s speech was made visible for reflection. As they transcribed their own speech, they noticed the gaps in their own linguistic knowledge. Fifth, the studies included opportunities for collaboration with a peer or peers, and teacher feedback. Working collaboratively to produce a transcript of their oral output with a peer created the opportunity to use language as a tool to mediate L2 knowledge, form and use, including discourse features. Because of the pedagogical focus, teacher feedback was included in the design of the studies, usually at the end of each transcription activity cycle. Another common factor was evaluation of the transcribing activity. All the studies compared the performances of the participants before and after the transcribing activity. As there were no specific language features targeted, the measures used to evaluate the performances varied. Overall, the studies showed that students made gains from the activity, which justifies the inclusion of transcription activities in the classroom.

The students’ feedback in all the studies was positive. They found it useful to transcribe themselves, and that the transcripts were a starting point for making improvements. Lynch (2007) and Mennim (2003, 2007) also found that gains made by the students were durable. They both acknowledged that a repetition effect could have accounted for the subsequent improvement in production. The variation in the outcome among the students was explained in a number of ways. In Mennim’s (2003) study, the different students’ focus was attributed to their different levels of proficiency in English. In Lynch’s (2001) study, the differences between one pair and the others was attributed to gender and cultural differences between the members of that pair, whereas in Lynch (2007), the difference in performance between the groups was attributed to
differences in the “depth of processing.” Taken together, the studies indicate the following: that self-transcribing is means for noticing spoken output; that long-term benefits accrue from noticing; and that students find this task to be beneficial.

While the studies are enlightening, there are some limitations. For example, there was little discussion about the process of noticing between the members of pairs, or the collaborative dialogue that went on between members of a pair. While Mennim (2012) discussed the strategies that were used to solve problems, little is said about how frequently these strategies were used for each type of error noticed.

A second limitation was that the question of whether there are differences between students who collaborate and noticed together, and those who notice alone was not addressed. In Lynch (2001, 2007) and Stillwell et al. (2010), the students produced the transcripts alone, but compared transcripts with a partner worked to produce a version together, that is, they collaborated to revise the transcript. In Mennim’s (2003, 2007, and 2012) studies, the students transcribed together, but each was responsible for making changes to his or her own part of the transcript. It was not clear from the studies what students noticed on their own, whether they could mediate their own learning, or how the presence of a peer mediated noticing.

A third limitation of the studies was that they failed to take into account student agency. Because the studies were conducted in the classroom, the procedures were designed to fit within normal class time and activity, and practical matters such whether the students had enough equipment to produce the transcription were considered. Lynch (2001) even considered whether the students would find the transcription activity boring. The researchers took into account student perception of the task by eliciting feedback from them about the activity; Stillwell et al. (2010) included the results of a survey of the students conducted at the end of the study. However, the feedback elicited was focused on whether the task was useful to the students. Little
was said about how the students felt about self-transcribing, whether they saw value in collaborating with others, and what effect this activity had on them.

And a final limitation is that there was little attention paid to the effects of noticing over the longer term. Only one study (Mennim, 2007) traced the effects of noticing over a longer period. But it was limited to just one pair of participants, and one language feature.

Thus, these studies laid the groundwork for my investigation. I decided to use transcription to initiate noticing of students’ oral output. While the transcription studies have found students to be capable of noticing a variety of errors, how much of each type was noticed, and whether noticing alone had different results that noticing with a peer remain unanswered. As mentioned earlier, the studies did not examine the collaborative dialogue that took place between pairs as they noticed/compared transcripts. Since there have been few longitudinal studies done in this area, I decided to design a study that took a longer term into account.

**Conclusion**

In this literature review, I traced the origins of the Noticing Hypothesis in brief, and discuss the general understanding of the role of attention in noticing. This includes a review of some of the research into noticing and attention from a cognitive perspective. Noticing is a well accepted concept in SLA and considered one of the main pillars of acquisition. I have noted that a Vygotskian-inspired lens offers an alternative view of learning as social in its origins, which enriches our understanding of language acquisition. As language itself is an important tool for mediating learning, the verbalization that is generated in collaborative dialogue studies as well as in stimulated recalls is a rich site for investigating noticing of language forms. In this review, I also discuss relevant studies that have used transcription to facilitate noticing in spoken output. The transcripts that were produced are artefacts that allow students to reflect further on their own
oral production. The conceptual frameworks and empirical work discussed coalesce and found the questions that drive my research. In this thesis, I aim to answer the following research questions.

**Research Questions**

Question 1: When given opportunities to notice, what do participants notice of their speech production?

Question 2: a) What do learners do with grammar, pronunciation, and content items when they notice them?

b) When given opportunities to repeat their speech performance, what do they do with the items they have noticed?

c) Do the items they have noticed become more accurate in their production over time?

Question 3: What do learners who work alone and those who work collaboratively do with what they notice?

Question 4: Is the verbalization during noticing, and its effects, different between learners who work alone and those who work collaboratively?

In the next chapter, I present and explain the methodology I used to collect the data and the procedures adopted for data analysis.
CHAPTER 3: METHODOLOGY

In this chapter, I explain how the study was conducted, what data were collected, and how the data were processed. I describe the study as longitudinal because I followed each participant over the course of an academic term and collected a vast amount of data for each one. In the first section of the chapter I describe the educational environment and the study participants. I also describe the examination in English fluency that the participants were preparing for as part of their preparation for meeting university entrance requirements. The initial data for this study were generated from the mock exam speaking practices that began the iterative processes used in this study. In the second section, I describe the research design: how the participants were organized, and what data were collected. In the final section, I explain how the data were organized and coded. I also explain which data were examined and analysed to answer the research questions, and how the data are presented in later chapters.

Section One: Background

Context of the Study

The data for this study were collected at a Band One secondary school for girls in Hong Kong. This school has a good reputation and is known both for its academic and its extra-curricular achievements. Students follow the curriculum prescribed by the Education Department, which leads to secondary level certification at Form 5, and matriculation at Form 7. At this school, all subjects, except Chinese and French, are taught in English. Among the

\[\text{Section One: Background}\]

\[\text{Context of the Study}\]

\[\text{The data for this study were collected at a Band One secondary school for girls in Hong Kong. This school has a good reputation and is known both for its academic and its extra-curricular achievements. Students follow the curriculum prescribed by the Education Department, which leads to secondary level certification at Form 5, and matriculation at Form 7.}\]

\[\text{At this school, all subjects, except Chinese and French, are taught in English. Among the}\]

\[\text{Hong Kong uses a banding system to describe student achievement at the primary school exit exams with Band One representing highest achievement and Band Three, the lowest. Thus, a Band One school would comprise of students with this level of achievement.}\]

\[\text{The Hong Kong education system has since changed such that students matriculate from secondary education after 6 instead of 7 years. Form 5 students are typically 16 – 17 years old, while Form 7 students are typically 18 – 19 years old.}\]
students, however, Cantonese is the preferred language of communication. As Li (2009) noted, using English in everyday conversation among Chinese Hongkongers is highly marked, except in the presence of non-Cantonese speakers.

The school system in Hong Kong is intensely competitive and access to higher education is based on an individual’s performance on the public examination. Tang and Biggs (1996) described the school system as “an exam-dominated environment” (p. 160), while Watkins (2010) stated that “the traditional emphasis on exams has been characteristic of the Hong Kong education system” (p. 72). The university entrance exams are high-stakes exams because they determine whether students can enter university and follow their preferred course of study, which for many determines their future career prospects and place in society.

Participants

My participants were students between the ages of 16 and 18, who were in Form 6 or 7. All of them had been at the school since Form 1. They were all local Chinese, and all of them claimed Cantonese as their mother tongue. They had all obtained, at minimum, an overall C in English on the Form 5 public examinations, which is equivalent to Band 5.92 - 6.40 on the IELTS scale.⁴ The teacher I approached when I was looking for participants for the study was the head of English at the school. At the time, she was teaching three senior English classes: 1 Sixth Form and 2 Seventh Form. She selected 8 students from each class to participate in the study. These were the students she thought would benefit from additional speaking practice in preparation for the speaking component of the university entrance examination. The name of the exam is Use of English.

---

⁵ IELTS refers to the International English Language Testing System. Overall Band 5 achievement is described as “Modest user,” while Band 6 is termed “Competent user.”
**Use of English: Paper D**

The *Use of English* examination aims to test the ability of candidates to understand and use English for practical communication (HKEAA, 2002, p. 1). It consists of five papers: listening, composition, grammar and usage, oral, and integrated skills. The oral component is Paper D, and it is administered in two parts. Part One is an individual presentation, and Part Two is a small-group discussion. Part One and Part Two are conducted at the same sitting. The examination is usually completed in 20 minutes.

For Part One, each candidate receives one of four thematically related prompt passages at the exam. The passages are about 300 words in length. Candidates are given 10 minutes to prepare to make an individual presentation. The instructions on the examination are as follows: “Read the following passage. You may make notes to use as the basis for your presentation, but should make the presentation in your own words. If you wish, you can add your own opinions or ideas.” At the end of the 10 minutes, four candidates, each with a different prompt passage (A, B, C, or D) are examined together. Figure 2 shows a typical seating configuration for the exam. The candidates sit with two examiners who grade each candidate independently.

*Figure 2. A typical seating arrangement for the Paper D examination.*
For Part One, the candidates present their passages in turn. Each candidate has 1½ to 2 minutes to speak. Candidates do not have the original passages in hand when they present, but they may use the notes they took while preparing to speak. Ideally, the candidates will listen carefully to each other’s presentations so that in the end all the candidates are aware of the information contained in all four prompt passages.

After all the passages have been presented, one of the examiners reads the instructions for Part Two, the small-group discussion, which is printed on all the prompt passages. For example, if the theme of the prompt passage is environmental protection, the candidates might be asked to discuss how, as a committee, they could organize activities for Environment Week in their school. This discussion is meant to simulate a study or work situation in which a problem or task needs to be addressed. The four candidates have 10 minutes for Part Two, during which they must demonstrate their ability to use English in a discussion. According to the syllabus (p. 3), the candidates are expected to demonstrate the following skills:

- express factual information, describing, explaining, confirming, presenting plans, etc.,
- express personal feelings and judgements, presenting arguments, agreeing and disagreeing, giving opinions, persuading, etc.,
- seek understanding and clarification through questioning and discussion,
- use appropriate interaction skills,
- pronounce English clearly and intelligibly, with the appropriate use of the intonation system.

The Paper D exam syllabus does not contain rubrics for scoring. Nor does it explicitly state that grammatical utterances are expected, though one would assume that in order to accurately and clearly convey meaning, sentences should be well-formed. In preparing students
for the exam, teachers rely on colleagues who have served as examiners to inform them about what is required of candidates. There appears to be no scale for what constitutes fluency, and as for content, there appears to be no yardstick indicating how much of the original content from the prompt passage needs to be covered, whether there are any deductions for misrepresentation of the content, or how much value is placed on a personal opinion given at the end of the presentation. There is also no rule about the voice used in the presentation: that is, about whether it is in the first person, or reported speech.

One of the strategies employed by students included memorizing the entire passage, although the instructions explicitly stated that the presentation was to be “in your own words.” Memorizing ensured the greatest chance of being grammatically correct, since all the sentences were produced verbatim from the prompt passage. A participant in the pilot study claimed that her examiners did not notice her presentation was a facsimile of the prompt passage, and happily revealed that this strategy earned her an “A” grade. It is not difficult to believe this claim. An examiner would hear at least 28 students on each examination day each presenting one of four passages, and could fail to recognize, especially at the start of the examination Session, a memorized prompt passage.

A sample set of exam papers showing four thematically related passages can be found in Appendix A.

Section Two: Research Design

Rationale for using Paper D

My study was planned around the Use of English Paper D exam for two reasons. First, the data collected would have ecological validity because it would be congruent with the curriculum, and the exam elicited the type of language produced by the participants when speaking in a formal context. The prompt materials were also from the collection used by the participants’
English teacher. Second, the students would be motivated to take part in the study because it offered them additional opportunities to practice for their exam.

**Pilot study**

In my pilot study, I used Part One of Paper D to elicit speaking data from the participants.

The research questions for the pilot study were as follows:

Question 1: Does transcribing assist learner noticing in oral language production?
Question 2: Does increased learner noticing result in increased grammatical accuracy in oral language production?
Question 3: Which produces greater noticing – when students transcribe in pairs, or alone?
Question 4: Under which condition is attention to the error most effectively directed: when the learners work alone, with a peer, or when given help from a teacher?

In brief, the pilot study found that the participants were able to transcribe their short presentations within a reasonably short time. As they transcribed, they were able to identify many of the mistakes that they had made, and make target-like corrections, which were used in the second attempt at the presentation.

**How the Research Questions and Instructional Intervention Evolved**

When the pilot study was completed, I reformulated the research questions for this study to take into account its longitudinal nature. I found that transcribing did help learners to notice, so instead of asking whether transcribing assisted learners in noticing, the research question evolved into RQ1, which asks what it was that participants noticed in their speech production.

RQ2 evolved to include specific Grammar, Pronunciation and Content items, and investigate the effects of repetition and durability over time. RQ3 was rewritten to explore the differences

---

6 A summary of the pilot study can be found in Appendix B
between individual learners and collaborative partners with regard to what was noticed, rather than focusing on which condition produced greater noticing. Similarly, RQ4 was revised because the research procedure no longer included teacher help, and explored in-depth the differences between the individual and collaborative conditions.

Evolving from the pilot study, the instructional intervention for this study operationalized noticing in three ways: self-transcription of the presentation; comparing the self-transcription with that done by a peer; and listening to the second and third attempt immediately after it is made. While there were only two opportunities for stimulated recall in the pilot study, a third opportunity was included in the main study. I elaborate further on the procedures used in the description of the data collection in a later section in this chapter.

**Organization of Participants**

The organization of the participants: by class; as units; and by condition – Individual or in a Pair; is shown in Figure 3. As previously mentioned, the 24 participants, 8 from each of three classes, were selected by their teacher. The participants from each class were divided into two Units, A and B, comprising 4 participants each, and remained in their units throughout the study. Once a week, the teacher would send either Unit A or B to me during an English period while the rest of the class continued with her. For example, in week 1, during the seventh period English class, Unit A would meet with me, and in week 2 during the same period, Unit B would meet with me. The units alternated weekly. When we met, we used a set of past papers to simulate a Paper D exam. The participants’ speaking performances were recorded. Over the course of the term, each Unit did six mock exam practices with me.

In addition to meeting as part of a Unit during class periods, each participant also met with me in a spare period or during lunch. I call these meetings “Sessions.” Each Session
involved participants meeting with me individually or in a pair. They were assigned to the
Individual or Pair condition based on their school schedule and availability, and remained in the
same condition throughout the study. Because Class 3 participants had no spare periods and
could only meet during lunch, all the participants in that class were assigned to work in Pairs.
This is illustrated in Figure 3. The participants from Class 1 and 2 were organized with an equal
number of Individuals and Pairs in each Unit, and they were assigned randomly to each
condition.

Figure 3. Organization of participants into units and assigned condition.
Data Collection Procedure: the Sessions

As mentioned previously, all the Part One Paper D mock exam presentations were recorded. These recordings formed the basis for the Sessions when participants met with me individually or as a pair outside of class time.

After each practice exam, the participants were sent a recording of a Part One presentation via e-mail. The participants were asked to transcribe the recording they were sent in the intervening week before they met with me in a Session. The recordings were between 1½ and 2 minutes in length. Participants in the Pair condition alternated between transcribing their own and their peer-partner’s presentation. Each participant in the Pair transcribed the assigned presentation on their own at home. Participants in the Individual condition alternated between transcribing their own file and the file of another in this condition. The participants were asked to transcribe exactly what they heard on the recording. At this point, all the participants had taken part in an introductory meeting at which I gave them instructions for preparing the transcripts. They were not required to use transcription symbols. They were asked to represent inaccurate or non-standard pronunciations in their own way, for example, “subtitue” for a mispronunciation of “subtitle.” Pauses were to be represented by long dashes or a series of periods, and dysfluencies, such as speech filler particles, were to be transcribed as well, for example, as “em,” “m” or “mm.” If they thought something was incorrect, they would indicate this by putting the correct form in parentheses next to the error. After transcription, the participants e-mailed the transcripts to me, which I printed in preparation for the Sessions.

When the participants met me for a Session, we would follow the stages illustrated in Figure 4. There were five stages in each Session, and the entire Session was recorded. In the Individual condition, the participant met with me during a spare period. At the meeting she was presented with two transcripts of her Paper D presentation: one that she had prepared and one
that had been prepared by another Individual participant, and asked to compare them. She was told to describe out loud the differences she observed between the two transcripts, and to mention anything else that she noticed or thought of while looking over the transcripts. She was encouraged to underline or mark up the transcripts to indicate the places where she noticed differences in the transcription. After this, the participant was given the original prompt passage and asked to do the oral presentation again. She was given 4 minutes\(^7\) to prepare. If anything was said during the preparation, it was considered part of the data generated at this stage, which is called Stimulated Recall 1 (SR 1). The data in SR 1 includes the transcription data, which had been done at home.

<table>
<thead>
<tr>
<th>SR 1 (includes transcribing at home)</th>
<th>PT 2</th>
<th>SR 2</th>
<th>PT 3</th>
<th>SR 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare transcripts + 4 minute preparation for PT 2</td>
<td>Listen to / comment on PT 2 + 4 minute preparation for PT 3</td>
<td>Listen to / comment on PT 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Stage 1 Stage 2 Stage 3 Stage 4 Stage 5

**Figure 4. The stages of a session.**

Following this, the participant presented the passage again. This second attempt is called Production Task 2 (PT 2), which is Stage 2 of the Session. The presentation was recorded on my computer.

\(^7\) As each class period was 45 minutes, in order to complete the Session in time for the participant to go on to her next class, the preparation time was kept short. If participants were unable to complete their preparation in 4 minutes, they were given a few more minutes, but most of the participants were able to complete their preparation in the allotted time.
The presentation was played back in the Session immediately after it was completed. The participant was encouraged to stop the playback at any time to comment or report on any thought that struck her as she listened. This is called Stimulated Recall 2, which is Stage 3 of the Session. After this, she was asked to do the presentation a third time, and again, given 4 minutes to prepare. If anything was said during this time, it was considered part of the data generated at Stage 3.

At Stage 4, the participant attempted the presentation a third time. This was also recorded on my computer. This third attempt is called Production Task 3 (PT 3).

Once again, the recording was immediately played back to the participant, and she was encouraged to stop the playback and comment on anything that struck her while she listened. This final stage is called Stimulated Recall 3 (SR 3).

If the participant was assigned to the Pair condition, she would follow the same stages in a Session. The only difference was the presence of a peer-partner who was invited to contribute her comments and observations during the Session. On average, each Session lasted 45 minutes, and every Session was transcribed.

Table 2 summarizes the data collection procedure described so far.

Table 2
A Summary of the Data Collection Procedure Cycle for One Session

<table>
<thead>
<tr>
<th>Time / Place</th>
<th>Who is involved</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1: Curriculum time</td>
<td>Units and researcher</td>
<td>Paper D practice: Part One and Part Two</td>
</tr>
<tr>
<td>/ in school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 1: Outside school hours</td>
<td>each participant</td>
<td>- Received a Paper D Part One recording</td>
</tr>
<tr>
<td>/ at home</td>
<td></td>
<td>- Transcribed the recording</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Sent researcher the transcript by e-mail</td>
</tr>
<tr>
<td>Week 2: Outside school hours</td>
<td>Individual participants + researcher</td>
<td>Session: SR 1, PT 2, SR 2, PT 3, SR 3</td>
</tr>
<tr>
<td>/ in school</td>
<td>Pair participants + researcher</td>
<td></td>
</tr>
</tbody>
</table>
As can be seen in Table 2, a typical cycle for collecting data was 2 weeks long. This cycle was repeated 6 times over the academic term. Thus, each participant transcribed six Paper D presentations over the course of the study: three presentations which they themselves had done, and three presentations done by another participant. Each participant was also the focus of a Session three times. For Individual participants, this meant that they met with me on three separate occasions, generating data for Session 1, Session 2 and Session 3. During each Session, the participants proceeded through all five stages described in Figure 4 above. For Pair participants, each in the Pair was likewise the focus of a Session three times. The difference between Individuals and Pairs was that in the Pair condition, the peer-partner was present in the Sessions and invited to contribute her ideas and make comments and suggestions. Thus, each participant in the Pair condition met with me in six Sessions, three in which they were the focus of the Session, and three in which they were the peer-partner.

In addition, I used two questionnaires in this study. The first was administered at the beginning of the study to collect background information regarding language use for each participant. The second questionnaire was completed at the end of the study to obtain feedback about the participants’ experiences in the study. These data were complemented by exit interviews conducted at the end of the study. The interviews allowed me to clarify some of the written responses, and gave the participants an opportunity to comment on their experience of the study beyond answering the structured questions that were posed in the questionnaires. This final meeting was also recorded and transcribed. The data were then organized and compiled for analysis. The participants were asked to choose a pseudonym for themselves to protect their identity. Some of their choices were rather unusual, but the pseudonyms reflected their personalities, and I use them to discuss the participants in the following chapters.

---

8 This questionnaire and a summary of the responses can be found in Appendix C and E.
9 The exit questionnaire can be found in Appendix D.
To summarize, the following data were collected from each student:

- Language background questionnaire
- 6 recorded Paper D presentations
- 6 transcripts of Paper D Part I (3 of self, 3 of peer)
- 3 recorded Sessions
- Exit questionnaire
- Exit interview

The study was limited to 24 participants in one school due to constraints of time and resources. Some difficulties encountered in the data collection process are described in Appendix K. In all, 68 Sessions were completed. The data from four Sessions were not included because the participants involved did not complete SR 3 (Stage 5).

Section Three: Data treatment

Organizing the Data

After reading through all the Session transcripts and the participant transcripts, the items that the participants noticed were identified. An item was considered noticed if the participant marked it on the transcript she produced or if it was mentioned during the Session. Sometimes, an item was both marked in the transcript and mentioned during the Session. At other times, an item was only marked without being mentioned, or only mentioned but not marked. Figure 5 shows examples of items that were marked on a transcript.

| Em Hello everyone, in this information and technology oriented age, other than cameras, mobile phones and smart cards, I think we are very familiar with using world wide web eh to search things in internet, right? And do you know how these search engines earn money? Eh it is by popping out advertisements. Em Anna |

Figure 5. Examples of items that were marked in a transcript produced by a participant.
I organized the items that were noticed into five categories: Grammar, Pronunciation, Content, General Comments, and Difference in Transcription. A detailed list describing how each item was categorized can be found in the Code Book. The data were entered into an 11-column grid to track the items over the course of each Session. To follow an item through the stages (columns 4 to 11), the columns were coded to summarize the information for the item. The codes contain information about when an item was noticed, what decisions (if any) were made, and if the subsequent attempt at the same item was successful. There were occasions when an item that was extensively discussed was either not re-used or the next attempt was only partially successful because participants had difficulty producing the correct pronunciation or remembering to use the correct verb. An example is given below in Table 3, with detailed explanations for each column shown in Table 4. This example is taken from a Pair Session in which the grammatical item that was noticed was the use of the singular versus plural form of the verb to be: “there is” versus “there are.” The item was marked in the peer-partner’s transcript, but there was no mention of it when the pair compared their transcripts. Later, at PT 2, the item was incorrectly produced again by the participant under focus. This time, at SR 2, the partner pointed this out to the participant under focus. As such, for the participant under focus, SR 2 was the Point of Noticing for this item, and the mistake was acknowledged and corrected. At PT 3, the same item was produced correctly, and at SR 3, the peer-partner noted that it was correct.

After this was completed, I abstracted from the grid information about when the item was noticed, whether the decisions and comments made about it were correct, if it was produced again, and whether the subsequent production was accurate. The information is summed up in 5 codes. In the example given in Table 3, the codes applied were -, I, C, C, C. The coding rules that I devised can also be found in Appendix I.

10 The Code Book and Coding rules are found in Appendix H and I.
Table 3
An Example of a Coded Item From a Pair Session

<table>
<thead>
<tr>
<th>#</th>
<th>C</th>
<th>Code</th>
<th>Transcript (SR 1)</th>
<th>Transcript Noticing (SR 1)</th>
<th>PT 2</th>
<th>Correct?</th>
<th>SR 2</th>
<th>PT 3</th>
<th>Correct?</th>
<th>SR 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>12</td>
<td>O-agree</td>
<td>there are many visible, both visible and hidden camera(cameras), em, in many place (places)</td>
<td>No mention</td>
<td>So there is actually ff- em, many visible and hidden camera around us.</td>
<td>No</td>
<td>C: there is many hidden cameras A: And then only that there is many hidden cameras C: yeah A: yeah R: what’s the correct version? C: there are many A: are many</td>
<td>Yes</td>
<td>C: because I think she have made less grammatical mistake actually in this time R: yeah? C: yes R: for example where? C: for example she, like she won’t use there is, which is following by something plural afterwards</td>
<td>Yes</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Column</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td># Number, Identifies the item. In the example given, this was the 75th item for this participant</td>
</tr>
<tr>
<td>2</td>
<td>C Code. The first number identifies the category. The second number identifies the Session – 1st, 2nd or 3rd. In the example, “12” refers to a Grammar item raised in the 2nd Session.</td>
</tr>
<tr>
<td>3</td>
<td>Code The first letter identifies the one who mentions the item first. The second set of information is related to the category of the item. In this example, this item was identified by the Other student in SR 2 (see below) regarding subject-verb agreement.</td>
</tr>
<tr>
<td>4</td>
<td>Transcript (SR 1) The transcript produced by the student(s) is shown. This is part of Stage 1, or SR 1. Here, the transcript produced by the peer-partner is shown. We see that the verb “are” was used correctly.</td>
</tr>
<tr>
<td>5</td>
<td>Transcript Noticing (SR 1) This contains excerpts of what was said about the item when the transcripts were compared. In this example, nothing was said because the verb “are” was used correctly.</td>
</tr>
<tr>
<td>6</td>
<td>PT 2 This is an excerpt from the Session transcript of the participant’s second attempt at the same presentation.</td>
</tr>
<tr>
<td>7</td>
<td>Correct? This tracks whether the item uttered in Production Task (PT) 2 is correct. In this example, the participant did not use the correct verb “are” before “many cameras”</td>
</tr>
<tr>
<td>8</td>
<td>SR 2 Stimulated Recall 2. This is excerpted from the Session transcript where the participants listen to and comment on the recording immediately after it is completed. Here, we see that Congee, the peer-partner, points out the mistake. She also supplied the correct answer, which Agatani, the participant under focus, repeated.</td>
</tr>
<tr>
<td>9</td>
<td>PT 3 This is excerpted from the Session transcript of the participant’s third attempt at the same presentation.</td>
</tr>
<tr>
<td>10</td>
<td>Correct? This tracks whether the utterance in Production Task (PT) 3 is correct. In this example, the participant has made the correction and used the correct verb “are” in relation to “many visible and hidden cameras.”</td>
</tr>
<tr>
<td>11</td>
<td>SR 3 SR 3 stands for Stimulated Recall 3. This is excerpted from the Session transcript at the point where the participants listen to and comment on the recording immediately after it is completed. In this example, Congee said that Agatani’s presentation contained fewer mistakes and mentions the successful correction of this item.</td>
</tr>
</tbody>
</table>
I have made every reasonable attempt to ensure that the coding is accurate and consistent, but a certain amount of interpretation was necessary. A difficulty in data processing was coding and judging the grammatical accuracy of the utterances. Unlike written data, spoken data are complicated by pronunciation, which makes it difficult to determine whether an error is due to mispronunciation or ignorance of a grammar rule. For example, when a participant used the verb “need,” but neglected to say the /-id/ ending, yet showed consistency in transcribing the same verb as “needed,” the question arose whether this spoken form is an error in grammar, or a mispronunciation. In some cases where it was consistent, I coded such items as pronunciation errors, particularly if the learner demonstrated a clear grasp of the rule involved in other sentences.

A random number generator (http://www.random.org/integers) was used to select 10% of the data, and a second rater, an experienced ESL teacher, was trained to code the selected data. The codes were compared and the inter-rater reliability was 86.3%. In discussion, the second rater and I were able to resolve most of the points of disagreement and arrive at a consensus. The five codes for each item were then entered into a statistical software package (SPSS version 17) and analyzed by categories, frequency of occurrence, time periods, and condition, that is, Individual or Pair. The coded data yielded 305 different patterns of code.

These codes were further organized in terms of “Point of Noticing” so that comparisons between categories, groups, and Sessions could be made. Point of Noticing is defined as the stage of the Session (see Figure 4 above) at which a participant first noticed an item. There were three possible stages at which an item might be noticed in each Session: at SR 1 (which includes transcribing at home, containing verbalized and non-verbalized data), SR 2, and SR 3.

Following this, the items were coded as to whether they were Judgement Possible (JP), or No-Judgement Possible (NJP). Table 3 above shows an example that was JP, that is, each step of
the process could be judged as Correct, Incorrect, Accept/No Change/ Incorrect, or Disappeared. Conversely, NJP items could not be labelled in the same way. These items comprised General Comments and Difference in Transcription items as well as a small proportion of items from the other categories.

To refine the analysis, items were sorted into “Trackable” and “Untrackable.” Untrackable items are those that cannot be tracked through the Sessions. Four types of items were considered to be Untrackable. The first type was Difference in Transcription because these were only mentioned at SR 1 and were not followed up. The second type was General Comment, because these were participant evaluations such as: “this time was better” and “still too fast.” A third type was items that were not used in PT 2 or PT 3 after being initially noticed at SR 1. Since they did not reappear, they were Untrackable. Lastly, items for which correct diagnostic remarks were made only at SR 3 were also Untrackable since this was the end of the Session.

Trackable items were those that could be followed through a Session. These items appeared at least twice: once at the Point of Noticing (First Noticing), and the second time, at the point of production (Product). Trackable noticing can only occur at SR 1 and SR 2. Items noticed at SR 3 are not trackable because they do not appear again after this stage. For some Trackable items, something that was said at SR 2 or SR 3 after the item was used again was labelled “Last Word.” Figure 6 illustrates how the three labels correspond to the coded data that were gathered from the Sessions.
Figure 6. Correspondence between coded data and trackable item labels First Noticing, Product, and Last Word.

An example of how the codes are labelled is shown in Table 5. In this example, First Noticing occurs at SR 2, when Aurora noticed that she had used the wrong form of a verb (catching) and corrected it. She reproduced the item at PT 3 where she correctly used “can catch.” This is the Product. Since she does not mention this item again at SR 3, there is no Last Word in this case.

Table 5
An example of First Noticing and Product Corresponding to Coded Data from Aurora’s Session 2, item #44

<table>
<thead>
<tr>
<th>SR1</th>
<th>PT2</th>
<th>SR2</th>
<th>PT3</th>
<th>SR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>...however using that ar card can help the police in catching and arresting the criminals</td>
<td>No mention</td>
<td>the police can arrest criminals and catching criminals</td>
<td>P: the police can arrest criminals, ... the police can arrest and catching so, uh delete the word catching. If I want to use the word ‘catching’ as a u- I should use ‘catch’</td>
<td>the police can arrest, criminals, uh arrests, criminals easier, um, ya. By checking the record inside the card, the criminals can catch the,唔係 the police can catch the criminals.</td>
</tr>
<tr>
<td>-</td>
<td>I</td>
<td>C</td>
<td>C</td>
<td>-</td>
</tr>
</tbody>
</table>

11唔係 is Cantonese, meaning “no.”
**Coding for Engagement**

To distinguish between the quantity and the quality of the spoken data generated by an item, I designed a 5-point code called “Engagement.” All Grammar, Pronunciation, and Content items were coded for this characteristic. Examples for each type of Engagement are found in Appendix J. “Cursory” (Type 1) Engagement characterizes items that were simply acknowledged, but where no further mention was made about what was noticed, or what correction was needed. When the participants said out loud what correction or change was needed, the item was coded as Type 2 or “Verbalized intended change.” When the participant used metalanguage or explained the reason for the change, the item was coded as Type 3, “Explanation / Metalanguage.” Type 4, “Unsure / question,” was used when a participant’s verbalization contained questions or showed that she was undecided about how to resolve a problem. Sometimes the solution was to refer to the original passage at the next opportunity during the Session. The last classification is “Only at SR 3” (Type 5), which was when the item was mentioned at the end of the Session. This has value for understanding what participants perceived of their performance, or what their thoughts at the end of the Session were. However, something mentioned only at the end of Session has limited use in illuminating how the participant arrived at PT 2 or PT 3.

By definition, General Comment and Difference in Transcription items were spoken data generated during the Session. Since Difference in Transcription items were essentially instances of cursory noticing without further follow-up, I did not consider these items in relation to Engagement. However, recurring themes and ideas mentioned in General Comments were taken into consideration. Items categorized as Grammar, Pronunciation, and Content were coded for Engagement. I found instances in which participants had noted an item when they transcribed the
first presentation, but did not mention it during the Session. I call these “non-verbalized” items. The rest were all considered verbalized items.

The results of the analysis for Engagement are presented in Chapter 5.

*Principles Used for Reporting Data*

In Chapters 4 and 5, the quantitative data are presented and discussed. I only comment on differences of 10% or greater. As the data were not normally distributed, non-parametric tests (chi-square) were used for the analyses. Analyses for which there are insufficient data are also not reported.

*Limitations of Quantitative Data*

There are limitations to the use of quantitative data. First, the quantitative data can give the impression that all items were noticed with equal attention, because each item is represented as one case in the dataset. For example, items that were noticed and corrected at SR 1 but not mentioned in the Session were counted as one case. However, items that were given considerable attention and discussed at length during the Session were also represented as one case in the dataset. To overcome this, I tried to show with the Judgement Possible, and Trackable variables that there were differences in what happened to an item after it had been noticed. However, a reading of the Session transcript remained the best way to get a sense of how much consideration was given to an item.

A second shortcoming is that the quantitative analysis does not account for items that were noticed, but remained unresolved, were avoided, or replaced. Participants often avoided the use of items that they were unsure how to improve despite having considered and discussed them. These items were counted as Untrackable because they were not used again, and thus, not
included in the quantitative results, masking the participants’ efforts and underrepresenting what participants noticed and engaged in.

A third limitation is that from the transcripts and self-reports of the participants, it was evident that participants had particular goals they were working towards, and often, these macro concerns were not captured in the Trackable items. Some examples of these concerns are: inclusion of particular discourse elements in the presentation; speaking slowly to eliminate dysfluencies; time management; maximizing fluency by taking few risks; and including an opinion. Both Individual and Pair participants were also observed to carry over ideas and new knowledge arising from previous Sessions into subsequent Sessions. All the goals that the participants focused on were difficult to quantify and discuss as single items in the quantitative database as most of these fell under the category of General Comments.

Although the quantitative analyses identified some differences in what was noticed between learners who worked alone and those who work collaboratively (RQ 3), a qualitative analysis of the Session transcripts was more illuminating in that it shed light on why Individuals and Pairs differed in what was noticed. The qualitative data are presented in Chapter 6.
CHAPTER 4: THE QUANTITATIVE DATA AND RESULTS

The results of the analysis of the quantitative data are presented in this chapter in three sections. The first section presents an overview of the dataset to address RQ 1. Section Two explains which data from the dataset are relevant to RQ 2a and RQ 2b, and uses those data to address those questions. In Section Three, I address RQ 2c, which asks whether noticed items become more accurate over time. I draw on the quantitative dataset using three different approaches to answer this question.

Section One

In this section, I address RQ 1 by reporting on what it was that participants noticed. RQ 1 is as follows: When given opportunities to notice, what do participants notice of their speech production? The participants noticed many things in their presentations: 2,174 items were noticed during the 68 recorded Sessions. For analysis each item was placed into one of five categories: Grammar, Pronunciation, Content, General Comments, or Difference in Transcription. Every category except the last contains a number of sub-items (types); these are listed in the Code Book in Appendix H. The largest number of items noticed was categorized under Grammar, followed by Content, General Comment, and Pronunciation. The smallest proportion, (175 items), were coded as Difference in Transcription (DIT). These proportions are illustrated in Figure 7. As DIT and General Comment items cannot be further analysed, these items will not be discussed further.

The number of observations made at each Session was similar. The number of items that were noticed at the first, second and third Sessions were 679 items (31.2%); 758 items (34.9%); and 737 items (33.9%) respectively. The number of items noticed over the three Sessions by the
participants ranged from 55 to 157. On average, each participant noticed $96^{12}$ items. The proportion of items noticed in each category did not change over the three Sessions.

![Distribution of coded items by category.](image)

**Figure 7. Distribution of coded items by category.**

**Grammar**

Grammar items comprised 45.9% of the total number of items noticed. On average, each participant noted 44 Grammar items over three Sessions, ranging from 14 to 82 items. Within this category, items were further divided by type, which can be seen in Table 6. The five most frequently occurring types of Grammar item noted were Subject-verb agreement (17.3%), Nouns (16.8%), Tense (15.3%), Sentence structure (12.3%), and Verb (9.0%). Together, these accounted for 70.7% of all Grammar items noticed. The remaining nine types listed in the table accounted for the remaining 29.3% of Grammar items. Verb-related Grammar items (Subject-verb agreement, Tense, Modal verbs and Verb) were noticed most frequently; as a group, these accounted for 45.5% of all items.

---

12 The calculation was based on 68 sessions. Thus, the total number of items (2,174) was divided by 68, and then multiplied by 3 sessions.
Table 6
Grammar Items Noticed by Type and Session

<table>
<thead>
<tr>
<th>Type</th>
<th>Session 1 (%)</th>
<th>Session 2 (%)</th>
<th>Session 3 (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject-verb agreement</td>
<td>50 (16.9)</td>
<td>51 (14.7)</td>
<td>72 (20.2)</td>
<td>173 (17.3)</td>
</tr>
<tr>
<td>Nouns</td>
<td>46 (15.5)</td>
<td>64 (18.5)</td>
<td>58 (16.3)</td>
<td>168 (16.8)</td>
</tr>
<tr>
<td>Tense</td>
<td>61 (20.6)</td>
<td>50 (14.5)</td>
<td>42 (11.8)</td>
<td>153 (15.3)</td>
</tr>
<tr>
<td>Sentence structure</td>
<td>34 (11.5)</td>
<td>43 (12.4)</td>
<td>46 (12.9)</td>
<td>123 (12.3)</td>
</tr>
<tr>
<td>Verb</td>
<td>23 (7.8)</td>
<td>31 (9.0)</td>
<td>36 (10.1)</td>
<td>90 (9.0)</td>
</tr>
<tr>
<td>Preposition</td>
<td>20 (6.8)</td>
<td>26 (7.5)</td>
<td>14 (3.9)</td>
<td>60 (6.0)</td>
</tr>
<tr>
<td>Article</td>
<td>20 (6.8)</td>
<td>22 (6.4)</td>
<td>15 (4.2)</td>
<td>57 (5.7)</td>
</tr>
<tr>
<td>Pronoun</td>
<td>9 (3.0)</td>
<td>22 (6.4)</td>
<td>19 (5.3)</td>
<td>50 (5.0)</td>
</tr>
<tr>
<td>Grammar comments</td>
<td>9 (3.0)</td>
<td>9 (2.6)</td>
<td>19 (5.3)</td>
<td>37 (3.7)</td>
</tr>
<tr>
<td>Modal verbs</td>
<td>11 (3.7)</td>
<td>8 (2.3)</td>
<td>18 (5.1)</td>
<td>37 (3.7)</td>
</tr>
<tr>
<td>Adverb</td>
<td>9 (3.0)</td>
<td>12 (3.5)</td>
<td>9 (2.5)</td>
<td>30 (3.0)</td>
</tr>
<tr>
<td>Comparatives / Superlatives</td>
<td>3 (1.0)</td>
<td>4 (1.2)</td>
<td>4 (1.1)</td>
<td>11 (1.1)</td>
</tr>
<tr>
<td>Phrasal verb</td>
<td>0 (0)</td>
<td>2 (0.6)</td>
<td>3 (0.8)</td>
<td>5 (0.5)</td>
</tr>
<tr>
<td>Gender</td>
<td>1 (0.3)</td>
<td>2 (0.6)</td>
<td>1 (0.3)</td>
<td>4 (0.4)</td>
</tr>
<tr>
<td>Total</td>
<td>296 (100.0)</td>
<td>346 (100.0)</td>
<td>356 (100.0)</td>
<td>998 (100.0)</td>
</tr>
</tbody>
</table>
Pronunciation

The participants noted 270 Pronunciation items, constituting 12.4% of all items. On average, each participant noticed 12 Pronunciation items, the range was 3 to 49 items, over the three Sessions. Four of the participants paid particular attention to pronunciation, and together, they accounted for 124 or 45.9% of all items in the category. Items in this category were coded under one of three types: Accept, Deny, or Comment on pronunciation. The greatest number were “Accept” items (65.6%). These items were those that the participants noticed and accepted as a pronunciation error. “Deny” items were those that participants noticed but considered to be mistakenly transcribed by the other participant, or not requiring action. These constituted about one-quarter (25.6%) of all Pronunciation items noticed.

Figure 8. Type of Pronunciation items identified by proportion.

Content

In this category, 413 items were observed, representing 19.0% of all coded items. On average, each participant noticed 18 items, the range was 8 to 41 items, over the three Sessions. One-third (n=8) of the participants accounted for half (n=208) of all items noticed in this category.
Items in Content were coded as one of five types. Items coded as Meaning/Lexis related to meaning at the level of the sentence affected by word choice. Items coded as Information addressed the absence, presence, accuracy, or choice of adding or omitting information in the presentation. Items relating to Opinion were comments about the absence or inclusion of a personal statement in a presentation. Discourse items related to the overall organization of information in the presentation and the order in which information was presented. Comment items were comments participants made about the content of the presentation in general. Figure 11 below shows the proportion of items coded by type. We see that Meaning/Lexis items were noticed most (43.6%), followed by Information items (28.3%).

![Figure 9. Type of Content items identified by proportion.](image)

In sum, to answer RQ 1, participants noticed many things (2,174 items) about their speech production. The largest proportion of noticed items came under the category Grammar, followed by Content and Pronunciation. Although the number of items noticed varied among
participants, no change was found over the three Sessions, and no differences were found between Individual and Pair participants.

Section Two

This section addresses RQ 2a and RQ 2b. The two questions are as follows:

2a) What do learners do with grammar, pronunciation and content items when they notice them?

2b) When given opportunities to repeat their speech performance, what do they do with the items they have noticed?

I begin with an explanation of which data were used to address these questions. This is followed by an analysis of the data to answer the research questions.

Point of Noticing

As mentioned in Chapter 3, Point of Noticing is defined as the stage in the Session at which a participant first notices an item. The majority of items (62.1%) were noticed during SR 1, followed by a drop to about 20.6% at SR 2, and further dropping to 17.3% at SR 3. Thus, the participants noticed more during SR 1 based on a transcript than they did by listening to a playback of the presentation at SR 2 or SR 3. Though participants were encouraged to review their performances during SR 2 and SR 3, few of them played back their recordings more than once. As a result, the participants did not detect as many items that required attention at SR 2 and SR 3 compared to SR 1. This occurred presumably because comparing transcripts at SR 1 used the visual mode and required more time than just listening to a playback, which encouraged greater reflection and engagement. The data also indicate that at SR 2 and SR 3 participants were
probably judging their own success in improving upon what they had noticed earlier rather than listening for new areas to improve.

The point at which something was noticed also depended largely on the category of the item. In Table 7, we see that more Grammar and Pronunciation items were identified at SR 1. Conversely, a sizable proportion of Content items relative to Grammar and Pronunciation items were noted at SR 2 and SR 3. A chi-square test was conducted to determine if Point of Noticing differed across categories of items, and the results were significant, $\chi^2 (4, n = 1,681) = 170.71, p = .001$. This is largely accounted for in Content, where point of noticing for items were tended to be more evenly distributed over a Session. Thus, we can see that comparing transcripts at SR 1 was helpful in drawing participants’ attention to Grammar and Pronunciation items.

**Table 7**

*Point of Noticing by Category*

<table>
<thead>
<tr>
<th></th>
<th>Grammar (%)</th>
<th>Pronunciation (%)</th>
<th>Content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SR 1</strong></td>
<td>761 (76.3)</td>
<td>203 (75.2)</td>
<td>171 (41.4)</td>
</tr>
<tr>
<td><strong>SR 2</strong></td>
<td>145 (14.5)</td>
<td>38 (14.1)</td>
<td>145 (35.1)</td>
</tr>
<tr>
<td><strong>SR 3</strong></td>
<td>92 (9.2)</td>
<td>29 (10.7)</td>
<td>97 (23.5)</td>
</tr>
<tr>
<td></td>
<td>998 (100.0)</td>
<td>270 (100.0)</td>
<td>413 (100.0)</td>
</tr>
</tbody>
</table>

**Judgement Possible (JP) and Trackable items**

As explained in Chapter 3, the data were organized into Judgement Possible (JP) and Non-Judgement Possible (NJP) items. In total, there were 674 NJP items, which constituted about 1/3 of all noticed items. These items offer insight into what was noticed, but because they cannot be further analysed, they are excluded from the following discussion.
On the other hand, JP items were those that contained diagnoses made by the participants. These could be judged as either correct or incorrect. There were 1,500 JP items. However, some of these items were not used again after they had been noticed. Thus, it was not possible to connect the item that had been noticed with a second or third production, that is, at PT 2 or PT 3. Hence, JP items were further categorized into Trackable and Untrackable items. Figure 10 illustrates the number of items in Grammar, Pronunciation, and Content that fall under each category.

Untrackable items, which cannot be followed through the Session, comprised 464 items or 21.3% of the total dataset. Since they could not be further analysed quantitatively, Untrackable items are also excluded from this discussion.
Conversely, Trackable items were items that were produced again in the Session after they had been noticed. In all, 1,036 items (47.6% of the total dataset) were Trackable. The analysis for RQ 2 is based on the data that fulfilled the criteria of being JP and Trackable.

To investigate RQ 2a (and 2b), it was necessary to match the two points of a Trackable item by First Noticing (FN) and Product. As explained in Chapter 3, these labels correspond to the codes that were applied to represent the stages of the Session. Based on an analysis for FN by category, the results related to RQ 2a are presented below.

Table 8 shows what learners did when they first noticed the JP items. In the majority of cases (81.2%), the items were noticed and changed so that they were correct. In contrast, a small number of identified items were misdiagnosed (7.1%), that is, the participants noticed a mistake and supplied an alternative that was also incorrect. The remaining JP and Trackable items were acknowledged, but the participants were unsure how to treat the items, or did not indicate what, if any, changes were to be made. To determine if items were first noticed and resolved differentially based on Category, a chi-square test was conducted. The results were significant, where \( \chi^2 (4, n = 1,036) = 98.41, p = .001 \). The differences are most striking for Pronunciation and Content items that were acknowledged without clear indication of what action was to be taken. Overall, it appears that participants noticed items that they could accurately resolve for themselves most of the time.
Table 8 shows that there were differences between categories of items. More Grammar items (85.8%) compared to Pronunciation (69.0%) and Content (76.8%) items, were first noticed and then correctly changed. The difference between Grammar and Pronunciation suggests that Pronunciation items were difficult to correct when they were first noticed. The other difference was that more Pronunciation and Content items than Grammar items were acknowledged without a clear indication of what the participant intended to change. Overall, the data suggest that Grammar items were more often successfully corrected after being noticed than Pronunciation and Content items.

Table 9 is a summary of Trackable items by FN, Product, and Last Word. If something was said after the item was first noticed and produced again, it would be labelled the “Last Word” about the item. Often, this was a confirmatory remark at SR 2 or SR 3 regarding the correct production of the item that had been noticed.
Table 9

Percentage of Trackable items at First Noticing, Product, and Last Word

<table>
<thead>
<tr>
<th></th>
<th>First Noticing</th>
<th>Product</th>
<th>Last Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct</td>
<td>842</td>
<td>802</td>
<td>169</td>
</tr>
<tr>
<td>Percentage</td>
<td>81.2%</td>
<td>77.4%</td>
<td>16.3%</td>
</tr>
<tr>
<td>Incorrect</td>
<td>74</td>
<td>234</td>
<td>6</td>
</tr>
<tr>
<td>Percentage</td>
<td>7.1%</td>
<td>22.6%</td>
<td>0.6%</td>
</tr>
<tr>
<td>Acknowledged, but no clear indication of action</td>
<td>121</td>
<td>-</td>
<td>27</td>
</tr>
<tr>
<td>Percentage</td>
<td>11.7%</td>
<td>-</td>
<td>2.6%</td>
</tr>
<tr>
<td>No last word</td>
<td>-</td>
<td>-</td>
<td>834</td>
</tr>
<tr>
<td>Percentage</td>
<td>-</td>
<td>-</td>
<td>80.5%</td>
</tr>
</tbody>
</table>

From the Product column in Table 9, it can be seen that after an item had been first noticed, a high proportion (77.4%) of the items were produced correctly at PT 2 or PT 3, while only 22.6% of the items were produced incorrectly.

From the Last Word column of Table 9, we see that it was common (80.5%) for participants not to mention an item again after it had been noticed and re-attempted. Of those who do mention the item, their comments were mainly correct (16.3%) as they affirmed that an accurate correction was made or that they had accurately resolved an outstanding problem. Conversely, the comments that participants made were seldom incorrect, although participants remained unsure of a small number (2.6%) of items. As only 19.0% of Trackable items have Last Word, I will not be discussing the results of analyses involving Last Word as a variable from this point.

**RQ 2b: When given opportunities to repeat their speech performance, what do learners do with the items they have noticed?**

To answer RQ 2b, the data for FN linked to the Product by Grammar, Pronunciation, and Content were analysed. From Table 10, it can be seen that among the 1,036 Trackable items, 65.5% of the items were first noticed and corrected, and then produced correctly again in the Session. Another 9% of items that were acknowledged at FN were also produced correctly later in the same Session. Together, these account for three-quarters (74.5%) of the items that were
produced correctly after FN. This suggests that noticing has a positive effect on subsequent production of the same item.

Table 10

Percentage of Items by Category Linking First Noticing with Product

<table>
<thead>
<tr>
<th>FN +</th>
<th>Product</th>
<th>Grammar (%)</th>
<th>Pronunciation (%)</th>
<th>Content (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct +</td>
<td>Correct</td>
<td>431 (65.7)</td>
<td>101 (58.4)</td>
<td>147 (71.0)</td>
<td>679 (65.5)</td>
</tr>
<tr>
<td>Correct +</td>
<td>Incorrect</td>
<td>132 (20.1)</td>
<td>19 (11.0)</td>
<td>12 (5.8)</td>
<td>163 (15.8)</td>
</tr>
<tr>
<td>Incorrect +</td>
<td>Correct</td>
<td>27 (4.1)</td>
<td>1 (0.6)</td>
<td>2 (1.0)</td>
<td>30 (2.9)</td>
</tr>
<tr>
<td>Incorrect +</td>
<td>Incorrect</td>
<td>35 (5.3)</td>
<td>4 (2.3)</td>
<td>4 (2.0)</td>
<td>43 (4.2)</td>
</tr>
<tr>
<td>Acknowledged, but no clear indication of action +</td>
<td>Correct</td>
<td>17 (2.6)</td>
<td>39 (22.5)</td>
<td>37 (17.9)</td>
<td>93 (9.0)</td>
</tr>
<tr>
<td>Acknowledged, but no clear indication of action +</td>
<td>Incorrect</td>
<td>14 (2.2)</td>
<td>9 (5.2)</td>
<td>5 (2.4)</td>
<td>28 (2.7)</td>
</tr>
</tbody>
</table>

|               |           | 656 (100.0) | 173 (100.0)       | 207 (100.0) | 1,036 (100.0) |

On the other hand, Table 10 also shows that 15.8% of items that were first noticed and corrected were produced incorrectly (row 2 of the table) later in the same Session. This suggests that some items continued to be incorrect even though they had been noticed and corrected.

The largest difference between categories (12.6%) was found between Content and Pronunciation items at FN(Correct) + Product(Correct). A difference can also be seen at FN(Correct) + Product(Incorrect), where Grammar and Content items differed by 14.3%. As well, when FN(Acknowledged) + Product(Correct) is considered, a large difference (19.9% and 15.3%) can be found between Grammar and the other two categories. Taken together, the data show that while Grammar accounted for the greatest proportion of FN(Correct) + Product (Correct) items, it was also the category with the most incorrectly produced items even if they
been corrected at FN. The data also indicate that FN had an important effect on Pronunciation and Content items: the number of incorrect products after FN were 18.5% for Pronunciation and 10.2% for Content, compared to 27.6% for Grammar.

In sum, in answer to RQ 2b, for JP Trackable items, the data indicate that participants were able to produce the majority (74.5%) of items again correctly if they had been corrected or acknowledged at FN. Content items were most often correctly produced after FN, followed by Pronunciation, and then Grammar items.

Section Three

RQ 2c: Are noticed items accurate over time?

I now turn to the final part of RQ 2. The answer to this question depends on how time is defined. As such, I approached the question in three ways using three different methods of analysis. Also, to determine accuracy over time, only items that were JP and Trackable (1,500 items) were considered.

To answer the question of whether noticed items were produced more accurately over time, I began with a consideration of the short term. The short term is defined as the duration of a Session. From the earlier discussion of RQ 2b, it can be seen that a large proportion of items (77.4%, from Table 9) were first noticed and then subsequently produced correctly within a given Session. Thus, three out of four noticed items remained correct in the short term.

To determine if items remained correct over a longer period, a different approach was used. This second approach was to look for a particular item across Sessions, by considering the Unit of time to be the duration of the school term. I call this the matching approach. I was aware that many of the items that were noticed earlier did not recur in subsequent Sessions, since each Session was based on a different prompt passage. To determine the extent to which noticed items
recurred, I only considered JP and Trackable Grammar and Pronunciation\(^\text{13}\) items that were correctly diagnosed and correctly reproduced within Sessions 1 and 2. Next, I searched for these same items in Sessions 2 and 3, and if they recurred, considered whether they had been accurately produced or not. This narrowed the dataset to 292 Grammar and 98 Pronunciation items.

From a total of 390 correctly diagnosed items that were subsequently produced correctly, I only located 44 items that recurred across Sessions. This represented a mere 11.3% of the small subset of data that were initially correctly diagnosed and correctly reproduced. The 44 items comprised 32 (72.7%) Grammar items and 12 (27.3%) Pronunciation items. What was most positive was that 42 of the 44 were correctly produced in a later Session after FN. As well, 13 (29.5%) of the 44 items that had been noticed again in the subsequent Sessions were re-corrected.

In analysing these cases, it appears that many of these items could be considered slips of the tongue at the first instance where each was noticed, and already in the participant’s understanding and repertoire. When the same mistake was made again at a later Session, the participant had little trouble identifying the problem. That the mistake recurred was likely not from lack of understanding but lack of practice, and a result of automaticity in articulation. The example below illustrates this.

**Example 1.** Honey – “daily lives” vs. “daily life”  
(Honey, 81, Session 1, SR 2, noun)  
**Honey’s transcript**  
And then she also m avoid to use English in m in in daily life.  
**PT 2** And (1.0) yup, and then (1.0) [yee] and she and she also æ also avoid to use English in her daily lives.  
**SR 2** H: I think I should say daily life  
**PT 3** She also avoid to use English in her daily life.

---

\(^{13}\) I did not consider Content items because these tended to be specific to the particular passage that was being presented.
Honey’s transcript
- He uses English in his daily lives, and he pays attention to every signs, notices, advertisements, and announcement in which is written in English.
- He will, for example, in daily lives he will take a notebook with him.

SR 1
H: and then the same correction we make is “daily live”. Before I said ‘daily lives’ and then we both corrected into ‘daily life’.

PT 2
He, he looks out of English instead of just doing the papers in his æ in æ papers. He uses English in his daily life.

SR 2
H: Daily life instead of daily lives.

PT 3
He uses English in his daily life.

On the other hand, there were items that suggest that getting it right at an earlier Session may have been a matter of luck rather than an understanding of a grammar rule. Consider Aurora’s difficulty with the modal “may” in the example below. Her strategy was to decide based on what “sounds right” to her.

Example 2. Aurora – modal form “may”+ “speak”
(Aurora, 5, Session 1, at SR 1 and SR 2)

Aurora’s transcript
She will never take an empty seat on the MTR that is next to a foreigner, even when she is really tired. “I’m afraid the person might ask me something,” she admitted.

PT 1
ah she was afraid to sit beside a foreigner, because she afraid that the foreigner may speak to her may spoke (speak) to her.

SR 1
A: And, ya may speak is correct. And, … no no, no no, I said may may may I correct already

PT 2
she wants to avoid having contact having any contact with that foreigner because the foreigner may has may spoke may speak to her and may ask her something,

SR 2
A: I don’t know use may or might, ya may is better but I, I don’t know the two the use of migh might might, might and may.

PT 3
because um, the foreigner mayed, spoke to may speak to her something,

In Session 2, many instances of the modal form “may” were used, and all of them correctly. No mention was made of the modal form. But at Session 3, a similar problem with the modal form arose again, although with a different verb, “say.”
Many teachers would say that memorization is not an effective way of learning, but it works well for John.

And, he has good, uh some people may say that, um memorization is not a good way to learn English.

Some people may say that memorization is not a good way to learn English but it is a good way to learn English for John.

Again, Aurora’s strategy for resolving this problem was to assess what “sounds right,” and we see her repeating possible choices in turn to test them out. This behaviour suggests that she had not fully grasped how modals, as a class of verbs, operate in a sentence, although she used them all correctly in Session 2. I believe the modal items she got right were like sets of discrete items correctly collocated or paired with the verb. Without an underlying understanding of the rule, Aurora was confused when she came upon a novel case in Session 3, which led to “re-noticing” the modal “may” + verb construction.

Thus, using the matching approach to answer RQ 2c was problematic and the analysis was inconclusive, not least because of the limited number of matches (44 items). Among these, it was heartening to note that more than two-thirds recurred correctly without the participant even mentioning or noticing them, suggesting that they were now part of the participant’s repertoire. On the other hand, it was difficult to determine, as illustrated above, if errors were merely slips of the tongue, even when they were first noticed, or if indeed, future accurate production...
represents clear evidence that the form was well understood. Without more data, it is hard to know whether what was noticed remained accurate over time.

The third approach to answering RQ 2c was to determine if grammatical categories that were noticed in Sessions 1 or 2 were no longer noticed in Sessions 2 or 3, indicating that they no longer posed difficulty for the participant. On inspection of the data, this was not the case. The data confirmed that the same type of items continued to be noticed in later Sessions, particularly Plural/singular noun, Subject-verb agreement, and Verb tense items. I did not observe any pattern such that what was noticed transferred to an entire class of Grammar items in subsequent Sessions. Being able to correctly notice and solve an agreement problem in a sentence was not generalized to cover all recurrences of this form in future utterances. That is, mistakes continued to be made from Session to Session. I could not determine whether the non-reappearance of a category of grammatical item meant that the participant had grasped the form and would be able to use it accurately in future.

In sum, the answer to RQ 2c depends on how time is defined. Within the span or cycle of one Session, from PT 1, to transcription and SR1, and then PT 2 and PT 3, most of the items that were noticed and correctly revised remained accurate in the learners’ production for the rest of the Session. If we define time as the span of all the Sessions, the answer to RQ 2c is inconclusive because the data are limited. It is encouraging that two-thirds of the items in a small 44 item dataset were reproduced accurately and without the participants’ explicit noticing. However, it is difficult to determine if re-noticed items were slips of the tongue or forms that remained problematic for the participants. Third, noticed items as representative of entire categories of grammatical items did not become more accurately produced over time, as items of the same form continued to be noticed in Sessions 2 and 3. The data suggest that noticed items were treated in isolation as a problem to be “fixed” within a particular sentence. What was noticed and
corrected seemed to contribute little to improvement in the use of the form in general. The same types of errors continued to be made and noticed in subsequent Sessions.

**Chapter Summary**

In this chapter, I present a quantitative analysis of the data to address the first two research questions of the thesis. In answer to RQ 1, participants noticed many things about their own speech production. Primarily, they noticed items in Grammar, followed by Content, then Pronunciation. The participants paid less attention to pronunciation despite the fact that they were doing a speaking task. There were no changes over time.

In Section 2, I address RQ 2a, which asks what learners do with Grammar, Pronunciation and Content items when they notice them. It can be seen from Point of Noticing that most of the items were noticed at SR 1, confirming that transcript comparison engaged participants more intensely in noticing than listening to the voice recording.

Data that fulfilled the JP and Trackable criteria pertaining to Grammar, Pronunciation, and Content items were examined. For Trackable items, the overall picture is a positive one from the perspective of learning. The majority (81.3%) of Trackable items were correctly diagnosed at First Noticing, while only a small number of items that were identified were misdiagnosed (7.1%). The data indicate that when learners were able to identify the mistakes they had made, they were also able to accurately correct them in most cases. Some changes were simple (usually morphological, involving one word) and mechanical, for example the addition or deletion of an ‘-s’ or ‘-es’ to a noun or an ‘-ed’ to a verb to accurately reflect the plural form or mark the past tense without involving other parts of speech in the sentence. However, rule application was not used consistently as a strategy because the learners’ knowledge was often incomplete. Even when prompted to think of equivalents to try to deduce the correct answer, participants did not
have the linguistic resources to do so. Instead, they often relied on what “sounds correct” or what was more intuitively appealing.

For RQ 2b, which asks what learners do with the items they have noticed when given opportunities to repeat their speech performance, the data indicate that participants were able to correctly produce the same item at the next opportunity for 77.4% of the items that had been noticed. When the item was tracked to link how it was treated at FN with the Product later in the Session, the data indicate that Content items were produced correctly at the highest rate, followed by Pronunciation, then Grammar. Most Grammar items that were first noticed and corrected were produced correctly later in the Session. However, the largest proportion of items that were first noticed correctly, but produced incorrectly later in the Session were also Grammar items. Pronunciation and Content had fewer such items.

Finally, in Section Three of the chapter, I address RQ 2c, which asks if the items that were noticed remained accurate over time. Three different approaches were used to answer RQ 2c. In the short term, over the course of a Session, participants were able to produce most of the items (slightly more than three-quarters) correctly in most cases if the item had been noticed earlier. However, due to the limited data available, it was not possible to assess whether the item would be accurately produced in the longer term. There were also no discernible patterns suggesting that what was noticed transferred to an entire class of Grammar items in subsequent Sessions. Thus, to answer RQ 2c, the data suggest that in the short term, items that were noticed were accurately produced. However, we do not know whether they remain accurate in the longer term.
CHAPTER 5: DIFFERENCES BETWEEN INDIVIDUALS AND PAIRS

This chapter addresses RQ 3 and RQ 4. Using the quantitative data, the chapter focuses on Individual and Pair differences. In Section One, I address RQ 3, which asks what learners who work alone do with what they noticed in comparison with those who work collaboratively. A report on how Individual participants and Pair participants differed when they noticed (within a Session, and across the three Sessions), and the outcome of their noticing is presented. The analysis centres on the JP and Trackable dataset, which consists of 1,036 Grammar, Pronunciation, and Content items.

In Section Two, I address RQ 4, which asks if the level of engagement during noticing results in different outcomes for learners who work alone in comparison with learners who work collaboratively. In Chapters 2 and 3, I refer to verbalization as a means to find insight into the mental processes of the participants as they solve problems. Through cognitive mediation, or languaging (Swain, 2006), the participants’ linguistic resources and the depth of engagement with an item can be observed. For Individual participants, talking about what they were thinking as they worked through their transcript comparisons in SR 1 (which includes transcribing) was private speech that they were making audible. As members of Pairs work together, each becomes aware of the gaps in her knowledge. Through dialogue, they may clarify what it is that they know and arrive at a solution to the problem. The section addressing RQ 4 compares and contrasts the differences between Individuals and Pairs in terms of what was verbalized, and the engagement in what was said.

Section One

Point of Noticing within a Session

From Chapter 4, Point of Noticing data show that most items were noticed at SR 1 (which includes transcribing), and fewer were noticed as the Session progressed. In order to find out if
there were differences between Individuals and Pairs as to when they first noticed an item across a Session, a chi-square test was conducted. The results were significant, $\chi^2(2, N = 2, 174) = 8.671$, $p = .013$. In order to locate where the differences between Individuals and Pairs lie, further chi-square tests were run for Point of Noticing by Session with regard to Category of item. For Grammar, we find that Individuals and Pairs were significantly different at Session 2 and Session 3. For Content, the groups were significantly different only at Sessions 3, while there were no significant differences at other sessions, nor were there differences at any session for Pronunciation. The results are shown in Table 11 below.

Table 11

<table>
<thead>
<tr>
<th>Session / Category</th>
<th>d. f.</th>
<th>n</th>
<th>Pearson chi-square</th>
<th>p value</th>
<th>Cramér’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 / Grammar</td>
<td>2</td>
<td>346</td>
<td>6.75</td>
<td>.034</td>
<td>.14</td>
</tr>
<tr>
<td>3 / Grammar</td>
<td>2</td>
<td>356</td>
<td>9.73</td>
<td>.008</td>
<td>.17</td>
</tr>
<tr>
<td>3 / Content</td>
<td>2</td>
<td>142</td>
<td>7.28</td>
<td>.026</td>
<td>.23</td>
</tr>
</tbody>
</table>

The graph in Figure 11 shows the percentage of items noticed at SR 1, SR 2, and SR 3 for Individuals and for Pairs. Pairs noticed a decreasing number of items over the course of the Session. The trend was downward and linear. In contrast, Individual participants noticed the least number of items at SR 2, followed by an increase in the number of items at SR 3. Thus, Individuals noticed proportionately more items at SR 3 than Pairs.
This result could partly be a result of the study design. Because Individuals did not have a partner to converse with during the Sessions, it was expected that they would notice fewer items than Pairs. Proportionally, however, I also expected Individuals to show a downward and linear trend similar to that shown by Pairs. This result suggests that Individuals noticed differently from Pairs. While Pairs appeared to be focused on what had been noticed and discussed before, and noticed fewer items at the end of the Session, Individuals continued to discover items that needed attention at the end of the Sessions.

First Noticing (FN)

From this point onwards, the data analysis only takes into account JP and Trackable data to account for what happened to the items in the Session.

In Table 12, differences between Individuals and Pairs can be seen. Overall, the proportion of items that were correctly first noticed by Pairs was greater than for Individuals. At
Session 1, a large difference (16.2%) can be seen for FN(Correct) between Individuals and Pairs. As well, the difference in FN(Noticed but no clear decision) was 10.5% between Individuals and Pairs at Session 1.

Table 12

**Percentage of Trackable items at First Noticing by Session**

<table>
<thead>
<tr>
<th>At First Noticing</th>
<th>Correct (%)</th>
<th>Incorrect (%)</th>
<th>No clear decision (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>103 (68.2)</td>
<td>17 (11.3)</td>
<td>31 (20.5)</td>
<td>151</td>
</tr>
<tr>
<td>Pairs</td>
<td>135 (84.4)</td>
<td>9 (5.6)</td>
<td>16 (10.0)</td>
<td>160</td>
</tr>
<tr>
<td>Session 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>120 (81.1)</td>
<td>10 (6.8)</td>
<td>18 (12.2)</td>
<td>148</td>
</tr>
<tr>
<td>Pairs</td>
<td>177 (82.3)</td>
<td>17 (7.9)</td>
<td>21 (9.8)</td>
<td>215</td>
</tr>
<tr>
<td>Session 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>119 (78.3)</td>
<td>12 (7.9)</td>
<td>21 (13.8)</td>
<td>152</td>
</tr>
<tr>
<td>Pairs</td>
<td>188 (89.5)</td>
<td>8 (3.8)</td>
<td>14 (6.7)</td>
<td>210</td>
</tr>
<tr>
<td>Trackables Only</td>
<td>842 (81.3)</td>
<td>73 (7.0)</td>
<td>121 (11.7)</td>
<td>1,036</td>
</tr>
</tbody>
</table>

At Session 3, the difference between Individuals and Pairs for FN(Correct) remained large at 11.2%. For Pairs, it was notable that at Session 3, First Noticing was almost 90% correct. While for both Individuals and Pairs the proportion of FN(Correct) items increased, the increase for Individuals from Session 1 to Session 3 was 10.1%, which is larger than the increase for Pairs.

This finding suggests that while both Individuals and Pairs benefited from the Sessions, Pairs were more likely to get the item correct at FN, and the proportion of items they corrected at FN increased from Session 1 to 3. For Individuals, although the proportion of items they correctly resolved at Session 3 was smaller than that of Pairs, the large increase from Session 1
to Session 2, and maintaining that at Session 3, suggests that they too had become more adept at identifying and correcting items over the Sessions.

A chi-square test was conducted to determine if Individuals and Pairs differed at First Noticing, and found significant differences between the groups, where $\chi^2 (2, n = 1,036) = 15.91$, $p = .001$. Statistically significant differences were found between Individuals and Pairs at Session 1, where $\chi^2 (2, n = 311) = 11.3$, $p = .004$; and Session 3, where $\chi^2 (2, n = 362) = 8.637$, $p = .013$. It can be seen that Individuals tended to be more indecisive at FN at Session 1, but with familiarity with the Sessions, the number of items that were first noticed with no clear decision decreased. Nonetheless, the difference between Individuals and Pairs persisted, and remained statistically significantly at Session 3, where we see Pairs more often correct at FN than Individuals. This supports the view that the peer-partner plays an important role in a Pair participant’s learning.

**Differences by Grammar, Content, and Pronunciation**

Table 13 shows the distribution of Trackable items for Individuals and Pairs by Grammar, Pronunciation, and Content. Over the Sessions, the proportion of Trackable items in each category changes.
Table 13
Percentage of Trackable items for each Category by Session

<table>
<thead>
<tr>
<th>At First Noticing</th>
<th>Grammar (%)</th>
<th>Pronunciation (%)</th>
<th>Content (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>91 (60.3)</td>
<td>35 (23.2)</td>
<td>25 (16.6)</td>
<td>151 (100.0)</td>
</tr>
<tr>
<td>Pairs</td>
<td>105 (65.6)</td>
<td>27 (16.9)</td>
<td>28 (17.5)</td>
<td>160 (100.0)</td>
</tr>
<tr>
<td>Session 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>79 (53.4)</td>
<td>32 (21.6)</td>
<td>37 (25.0)</td>
<td>148 (100.0)</td>
</tr>
<tr>
<td>Pairs</td>
<td>145 (67.4)</td>
<td>27 (12.6)</td>
<td>43 (20.0)</td>
<td>215 (100.0)</td>
</tr>
<tr>
<td>Session 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>95 (62.5)</td>
<td>31 (20.4)</td>
<td>26 (17.1)</td>
<td>152 (100.0)</td>
</tr>
<tr>
<td>Pairs</td>
<td>141 (67.1)</td>
<td>21 (10.0)</td>
<td>48 (22.9)</td>
<td>210 (100.0)</td>
</tr>
<tr>
<td>Trackables Only</td>
<td>656 (63.3)</td>
<td>173 (24.0)</td>
<td>207 (28.6)</td>
<td>1,036 (100.0)</td>
</tr>
</tbody>
</table>

With respect to Grammar, Pairs tended to identify about the same proportion of items at First Noticing for each Session. However, at Session 2, Pairs tended to identify proportionately more Grammar items than Individuals; the difference in proportion between Individuals and Pairs was 14.0%.

On the other hand, Individuals identified proportionally more Pronunciation items than Pairs. For Pairs, a steady decline in the proportion of Pronunciation items identified can be seen. At Session 3, the difference in proportion of items noticed by Individuals and Pairs for Pronunciation was 10.4%.

To find out if Individuals and Pairs differed in the number of Trackable items they identified in each category across Sessions, chi-square tests were conducted. Statistically significant differences were found between Individuals and Pairs at Session 2, where $\chi^2 (2, n = 363) = 8.23$, $p = .016$, and at Session 3, where $\chi^2 (2, n = 362) = 8.35$, $p = .015$. 
As expected, there was no significant difference between Individuals and Pairs at Session 1 because all the participants were familiarizing themselves with the procedure of the Sessions, and going about it in the same way. Over time, as they became more familiar with the procedure, differences between the groups emerged in what they focused on during each Session.

**Product**

From Table 14, we see that the number of items that were corrected after they were noticed increased proportionately over time for both groups. This suggests that both groups were becoming more accurate over time in reproducing items that they had noticed. The increase from Session 1 to Session 3 was 11.1% for Pairs. Correspondingly, the proportion of incorrect products produced by Pairs decreased from Session 1 to Session 3.

**Table 14**

*Trackable Correct, Incorrect, and Not Quite Correct Product by Session*

<table>
<thead>
<tr>
<th>Product</th>
<th>Correct (%)</th>
<th>Incorrect or Not quite correct (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>116 (76.8)</td>
<td>35 (23.2)</td>
<td>151</td>
</tr>
<tr>
<td>Pairs</td>
<td>114 (71.3)</td>
<td>46 (28.8)</td>
<td>160</td>
</tr>
<tr>
<td>Session 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>120 (81.1)</td>
<td>28 (18.9)</td>
<td>148</td>
</tr>
<tr>
<td>Pairs</td>
<td>158 (73.5)</td>
<td>57 (26.5)</td>
<td>215</td>
</tr>
<tr>
<td>Session 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>121 (79.6)</td>
<td>31 (20.4)</td>
<td>152</td>
</tr>
<tr>
<td>Pairs</td>
<td>173 (82.4)</td>
<td>37 (17.6)</td>
<td>210</td>
</tr>
<tr>
<td>Trackables Only</td>
<td>802 (77.4)</td>
<td>234 (22.6)</td>
<td>1,036</td>
</tr>
</tbody>
</table>

81
At Session 3, the proportion of incorrect products for Individuals remained high, while it fell 11.2% from Session 1 for Pairs. Thus, it appears that Pairs made greater gains than Individuals in correcting items after they had been noticed.

**First Noticing (FN) + Product by Session**

Following from Chapter 4, combining FN with Product can track an item after it was noticed. The proportion of items that Individuals and Pairs first noticed correctly and later reproduced correctly was similar, as were the other patterns for FN + Product.

Across the Sessions, however, we see that for both Individuals and Pairs the number of Trackable items that they correctly first noticed and correctly reproduced steadily increased, from 59.2% at Session 1 to 70.7% at Session 3. The difference is 11.5%. This suggests that both Individuals and Pairs became increasingly accurate in diagnosing what was noticed, and correctly applying this information in the following version of the presentation.

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (%)</th>
<th>Session 2 (%)</th>
<th>Session 3 (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct +</td>
<td>184 (59.2)</td>
<td>239 (65.8)</td>
<td>256 (70.7)</td>
<td>679 (65.5)</td>
</tr>
</tbody>
</table>

To answer RQ 3, there were differences between Individuals and Pairs in terms of what, when, and how they noticed mistakes/items. Individuals noticed more Pronunciation items than Pairs, and this behaviour persisted through the Sessions. Individuals also noticed more items at SR 3, suggesting that there remained unresolved items that continued to be noticed even at the end of the Sessions. In analysing the FN data, it could be seen the Individuals were more indecisive at FN, while Pairs tended to get more items correct at FN, suggesting that the presence of a peer-partner decreased indecision and helped to arrive at a correct change at FN. Similarly,
the presence of a peer-partner might explain why there were fewer items noticed at each successive SR. Perhaps the presence of a partner helped to maintain focus on what had been noticed at SR 1. The presence of a peer-partner could also account for the increasing proportion of correct products for Pairs from Session 1 to Session 3. Thus, we see that the Individuals and Pairs behaved differently with respect to what was noticed, when something was noticed, and how the item was treated after it was noticed. Table 16 below summarizes the salient findings.

Table 16  
Summary of salient findings for RQ 3

<table>
<thead>
<tr>
<th></th>
<th>Individuals</th>
<th>Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Point of Noticing</strong></td>
<td>More at SR3 than SR2</td>
<td>Least at SR3</td>
</tr>
<tr>
<td><strong>FN at Sessions</strong></td>
<td>At Session 1, bigger proportion of FN(Noticed, but no clear decision)</td>
<td>At Session 1, more FN(Correct) than Individuals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At Session 3, more FN(Correct) than Individuals</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increase in FN(Correct) from Session 1 to Session 3 is greater than Pairs.</td>
</tr>
<tr>
<td><strong>Identification of items</strong></td>
<td>Identified more Pronunciation items</td>
<td>Identified more Grammar items</td>
</tr>
<tr>
<td>(statistically significant difference found at Session 2 and Session 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Product</strong></td>
<td></td>
<td>Had greater proportion of correct products</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Had smaller proportion of incorrect products</td>
</tr>
<tr>
<td><strong>FN + Product</strong></td>
<td>Both Individuals and Pairs increased the proportions from Session 1 to Session 3</td>
<td></td>
</tr>
</tbody>
</table>

**Section Two**

**Counting What was Verbalized**

In this section, I address RQ 4, which asks whether the verbalization during noticing, and its effects, are different between learners who work alone and those who work collaboratively.
As mentioned in Chapter 3, the dataset was also coded for Engagement. Table 17 shows the number of verbalized and non-verbalized items in Grammar, Pronunciation, and Content.

Table 17

<table>
<thead>
<tr>
<th></th>
<th>Grammar (%)</th>
<th>Pronunciation (%)</th>
<th>Content (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbalized</strong></td>
<td>536 (59.0)</td>
<td>170 (18.7)</td>
<td>201 (22.2)</td>
<td>907 (87.5)</td>
</tr>
<tr>
<td><strong>Non-verbalized</strong></td>
<td>120 (12.0)</td>
<td>3 (1.1)</td>
<td>6 (1.6)</td>
<td>129 (12.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbalized</strong></td>
<td>656 (63.3)</td>
</tr>
<tr>
<td><strong>Non-verbalized</strong></td>
<td>173 (16.7)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>207 (20.0)</td>
</tr>
<tr>
<td></td>
<td>1,036 (100.0)</td>
</tr>
</tbody>
</table>

**Non-verbalized Items**

As can be seen in Table 17, non-verbalized items were primarily Grammar items. There are two possible reasons why this is the case. Firstly, it is easier to mark changes for Grammar when preparing a transcript of a presentation, whereas marking Pronunciation or Content requires more effort. For instance, visually representing erroneous pronunciation, or making an alternative suggestion for content requires more effort than correcting a verb form. Second, I suggest that task familiarity was an important factor. The English public examination contains a task that requires candidates to identify and correct grammar errors, one in each line, in a given passage. Since all the participants had considerable experience with this type of task in preparation for the public exams, it is unsurprising that this was their main focus when going over the transcripts.

An example is given below. In Example 3, we see a correction made at the transcription stage by the participant that was not discussed or verbalized during the Session, but was subsequently produced at PT 2 and PT 3 correctly.
**Example 3.** *(Buzz, 33, Session 2, SR 1, tense)*

Transcript: Have you ever try(tried) to eat organic food?
PT 2: Have you ever tried to eat organic food?
PT 3: Have you ever tried to eat organic food?

In Example 4, we see that Agatani’s peer-partner offered a correct suggestion on the transcript of Agatani’s presentation at SR 1. The item was not verbalized. At PT 2, we see that Agatani adopted her own incorrect suggestion, but at PT 3, she used her peer-partner’s suggestion with respect to one item, but failed to make the change suggested by her peer-partner with respect to another, which was to add “s” to the object “camera.”

**Example 4.** *(Agatani, 48, Session 2, SR 1, agreement)*

Own transcript: there are (is) visible camera to take down our movement
Other transcript: there are, are invisible camera [s] to take down our movements
PT 2: So there is actually ff- em, many visible and hidden camera around us.
PT 3: Yes there are many visible and hidden camera, around us.

Of the non-verbalized items in Grammar, Pronunciation, and Content, the majority (92.5%) were correctly diagnosed at the SR 1, but only 66.7% of these were correctly produced during the Session.

Table 18 shows the combined FN + Product proportions for verbalized and non-verbalized Trackable Grammar items. The proportion of items for both verbalized and non-verbalized items was similar. The majority of the items were correctly first noticed and correctly produced afterwards (Correct + Correct). However, the proportion of non-verbalized Grammar items that were correctly diagnosed and incorrectly produced (Correct + Incorrect) was 23.3%, almost 10% higher than that of verbalized Grammar items.
Table 18
Non-verbalized and Verbalized Grammar Items and Outcomes

<table>
<thead>
<tr>
<th>FN + Product</th>
<th>Non-verbalized (%)</th>
<th>Verbalized (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct + Correct</td>
<td>77 (64.2)</td>
<td>354 (66.0)</td>
<td>431 (65.7)</td>
</tr>
<tr>
<td>Correct + Incorrect</td>
<td>28 (23.3)</td>
<td>72 (13.4)</td>
<td>100 (15.2)</td>
</tr>
<tr>
<td>Correct + Not quite correct</td>
<td>6 (5.0)</td>
<td>26 (4.9)</td>
<td>32 (4.9)</td>
</tr>
<tr>
<td>All others</td>
<td>9 (7.5)</td>
<td>84 (15.7)</td>
<td>93 (14.2)</td>
</tr>
<tr>
<td>Total</td>
<td>120 (100.0)</td>
<td>536 (100.0)</td>
<td>656 (100.0)</td>
</tr>
</tbody>
</table>

There could be two reasons for the greater incidence of incorrect outcomes in spite of a correct diagnosis for non-verbalized items. First, since verbalization enhances memory, not mentioning the item during the Session meant that participants were more likely to forget the correction that had been made. Second, since verbalizing is a form of mediation, not verbalizing the item implies a lost opportunity to rehearse and solidify the improvement that had been noted on paper. Almost a quarter of the non-verbalized items that were initially correctly diagnosed were incorrectly produced the last time they were used.

**Verbalized Items**

The rest of this chapter is devoted to an analysis of the verbalized data. The data show that the category of an item determines to a large extent whether it is verbalized, and whether it is JP and Trackable. Pairs noticed most of the Content items. As well, Content items constitute the smallest proportion of Trackable items, and a substantial proportion (28.3%) of verbalized NJP items. This is perhaps due to the inherent nature of Content items. Participants’ discussions of what information to include, the right word to use, and how to organize the presentation tended to be difficult to judge. Languaging serves a macro function that controls and directs attention to
features of the presentation as discourse, such as the absence or presence of an introduction, conclusion, or opinion, and its overall organization. Because a partner was present, Pairs could verbalize and draw attention to these features in each other’s presentations. Individuals tended not to talk about the discourse features of their presentation.

**Engagement**

As well as distinguishing which items were verbalized and which not, I further coded the verbalized items by Engagement. Examples of each type of Engagement are found in Appendix J. Table 19 summarizes the number of items by category and Engagement. Most of the verbalization (47.9%) was Type 2 (see total column), that is, the participants said out loud what the correct version was or the change should be. Types 1, 4 and 5 characterize low levels of Engagement. Types 2 and 3 characterize high levels of Engagement.

### Table 19

**Grammar, Pronunciation, and Content items by Engagement**

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Grammar</th>
<th>Pronunciation</th>
<th>Content</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>Pair</td>
<td>Individual</td>
<td>Pair</td>
</tr>
<tr>
<td>1 Cursory</td>
<td>12 (3.6)</td>
<td>38 (8.0)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 Verbalized intended change</td>
<td>170 (50.3)</td>
<td>248 (52.2)</td>
<td>80 (58.0)</td>
<td>64 (52.0)</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>82 (24.3)</td>
<td>117 (24.6)</td>
<td>46 (33.3)</td>
<td>45 (36.6)</td>
</tr>
<tr>
<td>4 Unsure / question</td>
<td>16 (4.7)</td>
<td>24 (5.0)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5 Only at SR 3</td>
<td>58 (17.2)</td>
<td>48 (10.1)</td>
<td>12 (8.7)</td>
<td>14 (11.4)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>338 (100.0)</td>
<td>475 (100.0)</td>
<td>138 (100.0)</td>
<td>123 (100.0)</td>
</tr>
<tr>
<td>Total for each Category</td>
<td>813</td>
<td>261</td>
<td>404</td>
<td>1,478</td>
</tr>
</tbody>
</table>
To find out if there were differences in Engagement between Individuals and Pairs for Grammar, Pronunciation, and Content, a chi-square test was conducted. The results are presented in Table 20. We see that there is a statistically significant difference between Individuals and Pairs only with respect to Grammar items. With respect to Content items the difference approaches statistical significance. For Grammar, the results are largely accounted for by differences located at two types of Engagement – Type 1 and Type 5. As shown in Table 19, a larger proportion of items were cursorily engaged in (Type 1) by Pairs than by Individuals, while a smaller proportion of Grammar items engaged in were categorized as Type 5 (SR 3 only). However, as we see in Table 20, Cramér’s V value is small, so this statistic does not have much explanatory power. Nonetheless, it highlights the fact that Individuals defer verbalizing until the end of the Session, as is seen in the larger proportion of Content items of Type 5 Engagement in Table 19. This means that they had fewer opportunities to mediate their own learning during the Session, and missed the opportunity to test out the product of this mediation in PT 3.

Table 20
Results of Chi-square Tests Locating Significant Difference by Grammar, Pronunciation, and Content for Individuals and Pairs for Engagement

<table>
<thead>
<tr>
<th>Category</th>
<th>d. f.</th>
<th>n</th>
<th>Pearson chi-square</th>
<th>p value</th>
<th>Cramér’s V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammar</td>
<td>4</td>
<td>813</td>
<td>14.09</td>
<td>.007</td>
<td>.14</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>2</td>
<td>262</td>
<td>1.08</td>
<td>.582</td>
<td>.06</td>
</tr>
<tr>
<td>Content</td>
<td>3</td>
<td>404</td>
<td>7.46</td>
<td>.059</td>
<td>.13</td>
</tr>
</tbody>
</table>

Engagement over Session: Grammar, Pronunciation, and Content

Do Individuals and Pairs engage in items differently from Session to Session? To find out, Engagement data by category and Session were analysed. For Grammar (Table 19), the main difference was accounted for in the proportion of items that were verbalized and engaged in at
Type 5 Engagement. The proportion rose for Individuals at each successive Session, so that the greatest contrast to Pairs was at Session 3. With respect to Pronunciation, the largest differences were noted in Type 5 Engagement. Pairs engaged in proportionately more items at this level in Session 3.

Similarly for Content, the greatest divergence between Individuals and Pairs was seen at Session 3. This is illustrated in Figure 12. Again, Individuals engaged with proportionately more items for the first time at SR 3 (Type 5) for Content, diverging from Pairs who employed mostly Type 2 Engagement. Based on the three analyses, at Session 3, Individuals and Pairs engaged in Grammar, Pronunciation, and Content differently.

![Figure 12. Differences in Engagement in Content at Session 3 between Individuals and Pairs.](image)

In the following analysis, I only consider Judgement Possible and Trackable items, so that we can assess the outcome of noticing for these items. Because of the small number of items coded as Type 4 and 5 for Engagement, the items in these categories were combined and will
henceforth be known as Type 4 Engagement. The items will be analysed in relation to First Noticing (FN), their Product, and the two combined: First Noticing + Product.

**Engagement of Verbalized and Trackable Items by Grammar**

Table 21 shows the Engagement for Verbalized and Trackable items in Grammar at Session 3, where the differences between Individuals and Pairs were sharpest. The difference in proportion between Individuals and Pairs for Type 1, 2, and 3 Engagement ranges from 9.9% to 15.8%. The greatest difference was for Type 2 Engagement.

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Individual (%)</th>
<th>Pair (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Cursory</td>
<td>2 (2.6)</td>
<td>15 (12.5)</td>
</tr>
<tr>
<td>2 Verbalized intended change</td>
<td>52 (67.5)</td>
<td>62 (51.7)</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>15 (19.5)</td>
<td>35 (29.5)</td>
</tr>
<tr>
<td>4 Unsure / question + Only at SR 3</td>
<td>8 (10.4)</td>
<td>8 (6.7)</td>
</tr>
<tr>
<td>Total</td>
<td>77 (100.0)</td>
<td>120 (100.0)</td>
</tr>
</tbody>
</table>

**Engagement of Verbalized and Trackable Items by Pronunciation**

As discussed earlier, Pronunciation is the only category in which Individuals noticed more items than Pairs. Most of the items were self-identified for Individuals, but in Pronunciation, the majority of Verbalized and Trackable items were other-identified for both Individuals and Pairs. These items elicited verbalized responses from participants more frequently.
Individuals identified more Deny items than Pairs did. From Table 22, the difference between Individuals and Pairs is clear in Session 2 and Session 3. The difference in the proportion of Accept and Deny items ranged from 21% to 35%. For Pairs, there is a large decline so that by Session 3, only one Deny item was identified (or 4.8% of items). Meanwhile, for Accept items, we see a large increase in the proportion of items, from 59.3% to 95.2%, from Session 1 to 3. These data suggest that Pairs tended to converge over the Sessions in agreeing that an item had been mispronounced.

**Table 22**

*Pronunciation: Accept, Deny, and Pronunciation Items Identified by Individuals and Pairs Over the Sessions*

<table>
<thead>
<tr>
<th>Pronunciation category</th>
<th>Individual</th>
<th>Pair</th>
<th>Individual</th>
<th>Pair</th>
<th>Individual</th>
<th>Pair</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td>(%)</td>
<td></td>
</tr>
<tr>
<td>Accept</td>
<td>21 (60.0)</td>
<td>16 (59.3)</td>
<td>13 (41.9)</td>
<td>19 (73.1)</td>
<td>22 (71.0)</td>
<td>20 (95.2)</td>
<td>111 (65.3)</td>
</tr>
<tr>
<td>Deny</td>
<td>12 (34.3)</td>
<td>11 (40.7)</td>
<td>18 (58.1)</td>
<td>6 (23.1)</td>
<td>8 (25.8)</td>
<td>1 (4.8)</td>
<td>56 (32.9)</td>
</tr>
<tr>
<td>Neither accept nor deny</td>
<td>2 (5.7)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1 (3.8)</td>
<td>1 (3.2)</td>
<td>0 (0.0)</td>
<td>3 (1.8)</td>
</tr>
<tr>
<td>Total</td>
<td>35 (100.0)</td>
<td>27 (100.0)</td>
<td>31 (100.0)</td>
<td>26 (100.0)</td>
<td>31 (100.0)</td>
<td>21 (100.0)</td>
<td>170 (100.0)</td>
</tr>
</tbody>
</table>

In Table 23, the greatest differences between Individuals and Pairs involved Type 2 Engagement. For Individuals, most of the items identified were Deny items, while the reverse was true for Pairs, for whom most of the items identified were Accept items.
Table 23
Pronunciation: Verbalized and Trackable Items by Engagement for Individuals and Pairs

<table>
<thead>
<tr>
<th>Pronunciation category</th>
<th>Engagement</th>
<th>Type 2</th>
<th>Type 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbalized intended change</td>
<td>Explanation / Metalanguage</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
</tr>
<tr>
<td>Accept</td>
<td>26 (41.9)</td>
<td>29 (70.7)</td>
<td>29 (85.3)</td>
</tr>
<tr>
<td>Deny</td>
<td>33 (53.2)</td>
<td>11 (26.8)</td>
<td>5 (14.7)</td>
</tr>
<tr>
<td>Neither accept nor deny</td>
<td>3 (4.8)</td>
<td>1 (2.4)</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>62 (100.0)</td>
<td>41 (100.0)</td>
<td>34 (100.0)</td>
</tr>
</tbody>
</table>

Thus far, we have seen several differences between Individuals and Pairs with respect to Pronunciation. Individuals verbalize more Pronunciation items, and the majority of these are other-identified. For Individuals these are also mainly Deny items. It appears that when Individual participants mentioned Pronunciation items, they were reacting to other-identified items, usually disagreeing with the suggestion that the word had been mispronounced. A breakdown of these items by Engagement tells us that most of the Deny items were engaged as Type 2 items, without much elaboration. Over time, however, we find that for both Individuals and Pairs, the number of Deny items that were verbalized fell. Conversely, the proportion of Accept items rose for both Individuals and Pairs. We also find that Pairs engaged in about the same proportion for Accept and Deny items for Type 2 and Type 3 Engagement.

Engagement of Verbalized and Trackable Items by Content

Unlike Grammar and Pronunciation, the majority of Verbalized and Trackable Content items (50.2%) were Type 3 Engagement items while 43.8% were Type 2 items. This can be seen in Table 24. For Content, participants generated alternative versions that could better express their meaning, and offered explanations for why they were dissatisfied with the content. This usually involved Type 3 Engagement.
Once again, the greatest differences between Individuals and Pairs were found at Session 3. The difference between Individuals and Pairs was 19.3% for Type 2 Engagement, where Individuals engaged less than Pairs. Conversely, Individuals engaged at a higher proportion than Pairs at Type 3 Engagement, the difference being 23.8%. A difference of 10.4% in proportion is also found between Individuals and Pairs for Type 2 Engagement at Session 2.

Table 24  
Verbalized and Trackable Content Items Identified by Individuals and Pairs by Engagement Over Session

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Session 1</th>
<th></th>
<th>Session 2</th>
<th></th>
<th>Session 3</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
<td>Pair (%)</td>
</tr>
<tr>
<td>2 Verbalized intended change</td>
<td>7 (29.2)</td>
<td>7 (25.9)</td>
<td>20 (55.6)</td>
<td>19 (45.2)</td>
<td>9 (36.0)</td>
<td>26 (55.3)</td>
<td>88 (43.8)</td>
<td></td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>15 (62.5)</td>
<td>19 (70.4)</td>
<td>15 (41.7)</td>
<td>20 (47.6)</td>
<td>15 (60.0)</td>
<td>17 (36.2)</td>
<td>101 (50.2)</td>
<td></td>
</tr>
<tr>
<td>4 Unsure / question + Only at SR3</td>
<td>2 (8.3)</td>
<td>1 (3.7)</td>
<td>1 (2.8)</td>
<td>3 (7.1)</td>
<td>1 (4.0)</td>
<td>4 (8.5)</td>
<td>12 (6.0)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>24 (100.0)</td>
<td>27 (100.0)</td>
<td>36 (100.0)</td>
<td>42 (100.0)</td>
<td>25 (100.0)</td>
<td>47 (100.0)</td>
<td>201 (100.0)</td>
<td></td>
</tr>
</tbody>
</table>

Engagement of Verbalized and Trackable Items by FN

To consider Engagement in relation to what was first noticed, the data were compared for Grammar, Pronunciation, and Content by Engagement in terms of FN(Correct), which can be seen in Table 25. In general, a link can be made between Engagement and FN. Higher levels of Engagement (Type 2 and Type 3) were associated with a greater proportion of items that were FN (Correct) across all categories. For Pronunciation, Type 3 yielded the highest proportion of FN(Correct) results, but for Grammar and Content, Type 2 Engagement yielded the higher proportion of FN(Correct) items.
In Table 25, a difference in the proportion of Type 4 Engagement (11.9%) in Grammar between Individuals and Pairs can be seen. For Pronunciation, there were differences between Individuals and Pairs at both Type 2 (22.4%) and Type 3 (13.8%) Engagement. More Type 3 items were FN(Correct) than Type 2 items. For Content, Pairs had more items that were FN(Correct). In this case, however, the proportion of correct Type 2 Engagement items was higher than Type 3 Engagement items for both Individuals and Pairs. The data suggest that without a peer-partner, Individuals had proportionally fewer items (23.9% difference) correct at First Noticing for Content at Type 3 Engagement compared to Pairs. Taken together, the data suggest that the proportion of FN(Correct) was higher for Pairs at every level of Engagement.

Table 25
Engagement and FN(Correct) items for Grammar, Pronunciation, and Content

<table>
<thead>
<tr>
<th>Engagement</th>
<th>First Noticing</th>
<th>Grammar</th>
<th>Pronunciation</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
</tr>
<tr>
<td>1 Cursory</td>
<td>7 (70.0)</td>
<td>27 (77.1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2 Verbalized intended change</td>
<td>115 (86.5)</td>
<td>154 (89.0)</td>
<td>33 (53.2)</td>
<td>31 (75.6)</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>49 (86.0)</td>
<td>71 (89.9)</td>
<td>24 (70.6)</td>
<td>28 (84.4)</td>
</tr>
<tr>
<td>4 Unsure / question + Only at SR 3</td>
<td>11 (52.4)</td>
<td>18 (64.3)</td>
<td>1 (100.0)</td>
<td>-</td>
</tr>
</tbody>
</table>

Engagement of Verbalized and Trackable Items by Product

For Grammar, the proportions of correct Product for Individuals and Pairs were similar. Table 26 shows that the largest difference was found in Type 4 Engagement: Individuals and Pairs differed by 16.7% for both Correct and Incorrect Product. The data suggest that Pairs, if they were unsure, would make a mistake about half the time. It was surprising to find that Pairs had a higher proportion of incorrect Product than Individuals for Type 3 Engagement. One would
have expected discussion and verbal Engagement with a peer-partner to have yielded fewer incorrect products.

As for Pronunciation and Content, since the difference between Individuals and Pairs by Engagement for Product items after FN was less than 10%, these results will not be discussed.

Table 26
Grammar: Verbalized and Trackable Data Only by Engagement and Product

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Product</th>
<th>Correct</th>
<th>Incorrect / Not quite correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual (%)</td>
<td>Pair (%)</td>
<td>Individual (%)</td>
</tr>
<tr>
<td>1 Cursory</td>
<td>8 (80.0)</td>
<td>26 (74.3)</td>
<td>2 (20.0)</td>
</tr>
<tr>
<td>2 Verbalized intended change</td>
<td>99 (74.4)</td>
<td>136 (78.6)</td>
<td>34 (25.5)</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>42 (73.7)</td>
<td>56 (70.9)</td>
<td>15 (26.3)</td>
</tr>
<tr>
<td>4 Unsure / question</td>
<td>14 (66.7)</td>
<td>14 (50.0)</td>
<td>7 (33.3)</td>
</tr>
</tbody>
</table>

Engagement of Verbalized and Trackable Items by FN+Product

The data for Grammar FN+Product show that the Type of Engagement for Individuals and Pairs was proportionately similar for FN+Product. However, there were differences between Individuals and Pairs in Pronunciation and Content.

For Pronunciation, Table 27 shows that there were large differences for Type 2 Engagement (14.2%) between Individuals and Pairs for items that were FN(Correct) + Product (Correct). This suggests that the presence of a peer-partner was helpful. This can likewise be seen for Content items. In Table 27, we see that Type 3 Engagement items were more frequently correct for Pairs then for Individuals, and the difference was 18.6%, this pattern also suggests
that the presence of a peer-partner was helpful in both correcting an item at FN, and producing a correct Product later in the Session.

Table 27
Engagement and FN(Correct)+ Product (Correct) Items for Pronunciation and Content

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Pronunciation Individual (%)</th>
<th>Pair (%)</th>
<th>Content Individual (%)</th>
<th>Pair (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Verbalized intended change</td>
<td>29 (46.8)</td>
<td>25</td>
<td>27 (75.0)</td>
<td>43</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>21 (61.8)</td>
<td>23</td>
<td>27 (60.0)</td>
<td>44</td>
</tr>
<tr>
<td>4 Unsure / question</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Another area of difference between Individuals and Pairs can be found in items that were noticed but with no clear decision, which were later correctly produced. Table 28 shows these data for Pronunciation and Content. In the table, for Pronunciation, we see that Individuals had proportionally more Type 2 items (14.4%) than Pairs. As well, Individuals also had more Type 3 items (11.5%) than Pairs. This was also seen in Content, where Individuals again had a greater proportion of Type 3 items than Pairs (16.4%). The reason why proportionately more of such items occur for Individuals was probably due to the absence of a peer partner. As noted earlier, Individuals tended to be more indecisive at FN, but verbalizing at Type 2 or 3 Engagement may have helped Individuals to produce a correct Product later in the session.

Table 28
Engagement and FN(Noticed, but no Clear Decision)+ Product (Correct) Items for Pronunciation and Content

<table>
<thead>
<tr>
<th>Engagement</th>
<th>Pronunciation Individual (%)</th>
<th>Pair (%)</th>
<th>Content Individual (%)</th>
<th>Pair (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Verbalized intended change</td>
<td>21 (33.9)</td>
<td>8</td>
<td>5 (13.9)</td>
<td>4</td>
</tr>
<tr>
<td>3 Explanation / Metalanguage</td>
<td>7 (20.6)</td>
<td>3</td>
<td>13 (28.9)</td>
<td>7</td>
</tr>
<tr>
<td>4 Unsure / question</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

96
What accounts for the differences between Individuals and Pairs in Pronunciation? It seems curious that Individuals were more willing to discuss and work on Pronunciation during the Sessions than Pairs compared to other categories. I believe one important factor explains why Pronunciation was less attended to and discussed in Pairs: the presence of the peer-partner. Because a good relationship was crucial for working together, maintaining collegiality and face-saving accounts for much of the observed differences. In the absence of a peer-partner, Individual participants were at liberty to disagree with the way they were transcribed by the peer. Participants working in Pairs did not have this freedom since the peer-partner was present at the Session. Calling into question how something was transcribed can be confrontational because it suggests that the peer-partner failed to transcribe accurately or was too picky. It also puts the responsibility of being a model English speaker on the peer-partner. This can be a socially uncomfortable position for the partner. I had the sense that some peer-partner transcribers deliberately failed to highlight mispronunciations so as not to embarrass the participant under study. One participant noted in the exit interview her awareness of her partner’s lack of confidence in speaking, and said that it would be “mean” to identify the numerous mispronunciations because it would humiliate her partner. As Pairs identified fewer and fewer items over the Sessions, it can be seen that Pairs increasingly shifted their focus from Pronunciation to other aspects of the presentation that were less threatening to the collegiality of the relationship. We see evidence for this in the large number of Accept items identified by Session 3 while only one Deny item was identified (see Table 22). As well, the total number of items identified by Pairs fell to the lowest level in the third Session.

Thus, to answer RQ 4, it can be seen that verbalizing during noticing differed between Individuals and Pairs, and had some different effects as well. The data show that there was a greater incidence of non-verbalized items being rendered incorrectly even if they had been
corrected when they were first noticed. The data also show that Individuals and Pairs engaged in items differently depending on category. For Grammar, Individuals and Pairs differed in the frequency of Type 1 (Cursory) and Type 5 (engaging only at SR 3) Engagement. For Content and Pronunciation items, the difference in Engagement occurred when Individuals chose to engage in the item, usually at the end of the Session (Type 5 Engagement).

There were also differences found between Individuals and Pairs across the Sessions. The greatest difference was seen at Session 3. Pairs verbalized less about Grammar and Content, and proportionately more about Pronunciation at SR 3 in Session 3, while Individuals verbalized more about Grammar and Content at SR 3 in Session 3. Pairs also showed an interesting pattern of change in Engagement at Session 3. I contend that these changes result from a familiarity with the Sessions and the partners figuring out which types of verbalizing were most helpful.

In terms of Engagement, the majority of Verbalized and Trackable Grammar and Pronunciation items were Type 2 items, while for Content, the majority (50.2%) were Type 3 items.

For First Noticing, Pairs performed better than Individuals. The smallest difference was found at Type 2 and Type 3 Engagement for Grammar, while the differences for Pronunciation and Content were higher. For Content, the difference was as much as 23.9% by proportion for Type 3 Engagement. Thus, we see a definite advantage in having a peer-partner present who can supply the correction at First Noticing.

Comparing First Noticing + Product by Engagement for Verbalized and Trackable items in Grammar, Pronunciation, and Content, the highest proportion of items that were correct at First Noticing and correctly produced (Correct + Correct) were found at Type 2 Engagement. That participants were able to correct what they noticed without much discussion suggests that these items were known and in the participants’ repertoire, but perhaps not mastered or at least
not to the extent that they could be produced automatically. In comparing Type 2 Engagement, we find that there were proportionately more Correct + Correct items for Pairs than Individuals in all categories. For Pairs, across categories, the outcomes were best for Content (82.7%), then Grammar (75.1%), and lowest (61.0%) for Pronunciation. For Individuals, Correct + Correct items at Type 2 Engagement for Content were 75.0%, 68.4% for Grammar, and also lowest for Pronunciation at 46.8%.

The tables also showed that with respect to Pronunciation and Content for Individuals, a large proportion of items were first noticed without verbalizing a clear decision, but were later correctly produced. This figure is high at Type 2 Engagement for Pronunciation at 33.9%, and lower for Content at 13.9%. This suggests that noticing the item sufficed for Individuals to correct it at next production, implying that Individuals already knew the correct version, but did not automatically produce it at the first attempt. In that case, effort was required to remember the item in order to produce it correctly at the next opportunity.

For Type 3 Engagement, we see that Pairs had proportionately more Correct + Correct items than Individuals, except for Grammar, where Individuals had a higher proportion of such items. However, the proportion of Correct + Correct Type 3 Engagement items was lower than Type 2 Engagement items, while in the case of Pronunciation the reverse is true. It seems that engaging with Pronunciation items led to a better outcome, whereas this was not necessarily the case for Grammar or Content items.

Unlike the earlier set of data that included Non-trackable items, Engagement for Verbalized and Trackable items changed over time. For Grammar, among Individuals, we see an increasing proportion of items engaged at Type 2 while the proportion of Type 3 items decreased. A similar trend was observed in Content for Pairs, where items were increasingly engaged at Type 2, and decreasingly engaged at Type 3. Also, we see that for Pairs,
Pronunciation was increasingly engaged with at Type 3 while items engaged at Type 2 decreased. While both types of Engagement correlate with a high proportion of Correct + Correct outcomes, I believe that Type 3 Engagement decreased over time because Pairs and Individuals came to realize that it was more efficient to focus on items that could be easily repaired and immediately helpful for the next speaking task than to spend time negotiating and discussing items that they found difficult to resolve.

In general, thus, we see that items that were verbalized at First Noticing and engaged actively (either Type 2 or Type 3) were often correct in the subsequent attempt, depending on the category of the item. Table 29 below summarizes the salient results.
Table 29  
*Summary of salient findings for RQ 4*

<table>
<thead>
<tr>
<th>Non-verbalized + verbalized Grammar items</th>
<th>Individuals</th>
<th>Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>For both Individuals and Pairs, fewer verbalized items were Correct + Incorrect, compared to non-verbalized items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbalized items</td>
<td>Noticed more Content items</td>
<td></td>
</tr>
<tr>
<td>Engagement (Differences only found in Grammar)</td>
<td>More Type 5</td>
<td>More Type 1</td>
</tr>
<tr>
<td>Comparing Engagement over Sessions</td>
<td>For Grammar, more Type 5 at Session 3</td>
<td>For Pronunciation, more Type 5 at Session 3</td>
</tr>
<tr>
<td></td>
<td>For Content, more Type 5 at Session 3</td>
<td>For Content, more Type 2 at Session 3</td>
</tr>
</tbody>
</table>

| At Session 3, Engagement by Grammar | | |
| Pronunciation | More Type 2 Deny items | More Type 2 Accept |
| Content | More Type 3 | |

| Engagement by FN(Correct) Grammar | | |
| Pronunciation | More Type 4 | |
| Content | More Type 2 and Type 3 | |

| Engagement by Product Grammar | | |
| Pronunciation | More Type 4 Correct, Fewer Type 4 Incorrect | |

| Engagement by FN(Correct) + Product (Correct) Pronunciation | More Type 2 | |
| Content | More Type 3 | |

| Engagement by FN(Noticed, but no clear decision) + Product (Correct) Pronunciation | More Type 2 | |
| Content | More Type 3 | |

---

**Chapter Summary**

In this chapter, the quantitative data are analysed and differences are found between Individuals and Pairs regarding when they noticed an item, what happened to that item when it was noticed, if it was used again later in the Session, and whether it was accurately produced.
Beginning with Point of Noticing, the data show that within a Session, there were significant differences between Individuals and Pairs. While there were no differences between Individuals and Pairs for Category of Trackable items noticed at Session 1, statistically significant differences were found both at Session 2 and Session 3, where Individuals noticed more Pronunciation Trackable items than Pairs did.

For First Noticing, statistically significant differences between Individuals and Pairs were also found, but only at Session 1 and Session 3, at which Pairs diagnosed proportionately more FN(Correct) items. Over time, we see that this proportion of items rose for both groups, while FN(Incorrect) and FN(Noticed, but no clear decision) decreased for both groups. The data suggest that both groups of participants improved at correctly resolving items that they had noticed over time, with Individuals making greater gains.

No statistically significant differences were found between Individuals and Pairs for Product, and by Sessions. Like First Noticing, both groups became increasingly adept at reproducing what they had first noticed over time, with Pairs making greater gains proportionately over time. Over the Sessions, we see increasing accuracy for both Individuals and Pairs for FN(Correct) + Product(Correct). Thus, while there were differences between Individuals and Pairs, both groups improved in terms of increasing the proportion of items they correctly diagnosed, and reproduced within the Session.

The data also show that verbalizing an item had a positive effect for later production. Non-verbalized items comprised a small proportion (9.4%) of all the items noticed. Of these items, many of them appeared to be mistakes that were correctly first noticed by the participants in their transcripts, but were not mentioned or discussed during the Session. The proportion of non-verbalized items that were FN(Correct) + Product (Correct) was comparable to the proportion of those that were verbalized. However, about one-quarter of non-verbalized items
which were correctly first noticed were incorrectly produced at PT 2 or PT 3, a proportion that is greater than for verbalized items. This suggests that when an item is verbalized, Engagement increases. Discussion or talk about the item tended to result in correction, while those that were not verbalized tended to be forgotten or incorrectly reproduced later in the Session.

As for differences between Individuals and Pairs in terms of how they verbalized and engaged with each item in relation to its outcome, the data show that Pairs have higher proportions of FN(Correct) + Product(Correct) outcomes compared to Individuals. The Engagement of Individuals and Pairs differs the most at Session 3 for all three categories. Type 2 Engagement in Pronunciation and Type 3 Engagement in Content produced the highest proportion of FN(Correct) + Product(Correct) items. In general, the data show that greater Engagement is associated with a higher proportion of Product(Correct) outcomes for noticed items, and that the presence of a peer-partner in Pairs had a positive effect on First Noticing as well as on the Product.

For mispronunciations, the participants’ difficulties were due primarily to lack of knowledge. Participants who worked alone only had an indication of the problem on the transcript, and they did not have access to a source of correct pronunciation during the Session. Members of Pairs could sometimes rely on the partner for help, but there were times when neither knew the correct pronunciation. Another difficulty was the physical articulation of sounds. Chan and Li (2000) and Chan (2006, 2007, 2009, 2010) have documented the difficulties faced by Cantonese ESL learners in pronouncing certain English sounds, which I also observed in my participants. The participants had difficulty with certain consonant clusters, such as in “access,” “texts,” and “relaxed.” Some also had difficulty producing words starting with /l/ and replaced it with /n/ at times, and vice-versa. For example, “like” was pronounced as “nike,” while “noon” came out “loon.”
In the next chapter, the qualitative data are used to deepen our understanding of verbalization and Engagement in Individuals and Pairs and the effect this had on how the participants approached their performances.
Having investigated the effects of engagement from a quantitative perspective, I now turn to a qualitative analysis of the data to address RQ 4, which asks whether engagement and its effects differ between learners who work alone and those who work collaboratively. The qualitative data enrich our understanding of the nature of engagement between the participants. I state in Chapter 3 some limitations of the quantitative dataset. A consideration of the qualitative data can fill this gap. Untrackable and no-judgement-possible items (General Comments and DIT items) account for a large proportion of the data. Though these items cannot be tracked over a Session, they offer information about participants’ noticing.

This chapter is divided into three sections. In the first section, I illustrate how Individual and Pair noticing were similar or different by using extended examples.¹⁴ In the second section, I highlight other differences not illustrated in the extended examples. In Section Three, I identify some of the macro-preoccupations of Pairs that developed through the Sessions to show how collaborative pair work gave rise to noticing and effects that were different from the effects of working on one’s own.

Section One

Differences between Individual and Pair engagement were expected as is implicit in the research design, but as Table 30 shows, the recording times for the two groups were similar. Individual recordings were slightly longer on average, and there was also a greater range among Individuals. However, Pairs said more, as seen in the average number of words and pages per Session transcript.

¹⁴ The data for these examples come from the Individual and Pair Sessions, language background questionnaires, end of study questionnaires, and exit interviews. I am using the term “extended examples” to characterise the descriptions of Individuals and Pairs, rather than using “case studies,” because this study was not conceived using case-study methodology (Duff, 2008).
Table 30  
**Individual and Pair Session Recording Time and Length of Transcripts**

<table>
<thead>
<tr>
<th></th>
<th>Individuals</th>
<th>Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Session (minutes: seconds)</td>
<td>45:13</td>
<td>44:41</td>
</tr>
<tr>
<td>Shortest Session (minutes: seconds)</td>
<td>26:15</td>
<td>30:55</td>
</tr>
<tr>
<td>Longest Session (minutes: seconds)</td>
<td>68:16</td>
<td>65:00</td>
</tr>
<tr>
<td>Average number of transcript pages / Session</td>
<td>7.6</td>
<td>10.8</td>
</tr>
<tr>
<td>Average number of words per transcript / Session</td>
<td>2,762</td>
<td>3,516</td>
</tr>
</tbody>
</table>

**Extended Examples**

To illustrate how verbalizing during noticing and its effects are different between Individuals and Pairs in Grammar and Pronunciation, I use extended examples from the data collected for 6 of the participants. Two Individual and 2 Pairs of participants were selected because they identified the greatest number of Grammar or Pronunciation items Individually or as a Pair in the study, as shown in Table 31.

Table 31  
**Number of Items Noticed by Extended Example Participants**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Number of items identified in 3 Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aurora (Individual)</td>
<td>82</td>
</tr>
<tr>
<td>Candy &amp; Ella (Pair)</td>
<td>46</td>
</tr>
<tr>
<td>Bobo (Individual)</td>
<td>49</td>
</tr>
<tr>
<td>TC &amp; Swatch (Pair)</td>
<td>26</td>
</tr>
</tbody>
</table>

**Grammar Extended Examples: Aurora (Individual), Candy and Ella (Pair)**

Aurora was a F.6 student who was assigned to the Individual condition. She was lively and motivated, and she approached the study with great enthusiasm. Unlike some of the other participants who worked alone, she showed little inhibition in talking out loud during SR 1 in her Sessions. Thus, her Sessions were some of the longest among the participants.
In her language background questionnaire, Aurora indicated that she liked English. She spoke English with her eight-year-old brother to encourage him to use it. Sometimes, she and her friends would pretend to be foreigners who spoke English with “very strong ‘r’ sounds.” Aurora liked practicing for the English oral exams. However, she lamented that “I don’t know why I always make so many grammatical mistake when I speak” and thought it was because “I speak too fast and the way I think is too slow.” She was confident that she could use English for a range of purposes, but it depended on whom she was talking to. If she perceived her interlocutor to be superior in status, she tended to lose confidence because “I don’t want to let them know that I have a lot of grammatical mistakes.”

On the exit questionnaire, in response to the question about what she had learned from transcribing herself, Aurora wrote: “I learnt to pay more attention on my tenses and other gramma.” She also wrote that the three things she learned from noticing her own speech was to “avoid starting a sentence with ‘and,’ speak slower, and think faster.” During the exit interview, she returned several times to her resolve to focus on grammar. She recognized that speaking was different from writing because there was no opportunity to review and edit while speaking. She thought she was compensating for her lack of grammatical accuracy while speaking by modulating her voice and pitch. Having had a positive experience at the mock exam the day before the interview, she told me she could “use my voice level and the up and down of my voice to attract people, and this help to score marks and ya, attract people. And they will ignore all my mistake.”

Given a choice, Aurora indicated that she would have preferred working with a partner. She thought that a peer-partner could have introduced an element of competition which could make her “improve faster.” She had heard from friends who were in Pairs that they had a lot of
fun working with their peer-partners. At the same time, she thought that if there had been another participant with her, she might have been “more relaxed, can deliver speech to her.”

During the Sessions, Aurora was able to mediate her own learning by listening to her own recording, and by testing out alternative versions. Below is an example in which she changed the position of the adverb ‘always’ and made a mistake in the form of the verb “to have.” However, she caught the mistake when she heard herself at SR 3 and corrected the verb.

**Example 5.** (Aurora, 40, Session 2, SR 2, agree)\(^{15}\)

SR 2 while listening to PT 2 “She has a smart card with her always,” Aurora stopped the replay
A: she has a smart card with her always. She always have a smart card with her.

SR 3 while listening to PT 3 “She always have her smart card”
A: Always has

One of the strategies Aurora tended to use to solve a problem was to “sound things out.” This was successful sometimes, while at other times it did not yield a solution or she ended up avoiding the item. Although she verbalized and tested out different versions, she was usually uncertain as to which was correct.

Another way of managing a difficulty she employed was avoidance. In Session 2, at SR 2, while listening to the recording of PT 2, Aurora stopped the replay at “If the card is being taken is getting into the wrong hand.” She asked: “is being taken, is taken, is being taken. Which one is better? Is being, is being taken.” Later at PT 3, the verb form for “be + take” was avoided and she said: “If the card is get into the wrong hands, for example.”

---

\(^{15}\) Each of these examples are labelled with the following information in parentheses: name of participant; the number of the item in the coded data set; Session 1, 2, or 3; stage at which the item was First Noticed (SR 1, PT 2, SR 2, PT 3, SR 3); and the type of item. Thus, for Example 6, this item was produced by Aurora; item 40 of her coded dataset; from Session 2; First Noticed at SR 2; and is an item on subject-verb agreement.
In an example where Aurora was deciding between two prepositions, “in” and “at,” it appeared that correct usage was sometimes a matter of chance. Earlier at SR 2, she made a decision to use “at” in the attempt, but at PT 3, she used “in.” Her final comment at SR 3 is shown:

A: And I use “at school,” “on MTR” even though I’m not so sure, but ya, just just all come out and then wow, no one will heard the mistake.

Example 6 below shows Aurora verbalizing various forms of “meet” that were appropriate for what she was trying to express. This example suggests that Aurora did not arrive at an understanding of when to use the appropriate tense or agreement. Instead, she was verbalizing in order to collocate appropriate verbs with “vocabulary” or “word,” and make a decision based on what sounded familiar.

**Example 6. (Aurora, 67 & 69, Session 2, SR 2, agree)**

PT 2 so, when he met when he meet some new vocabulary

SR 2
A: when he meets ya I want to say. Met is the past tense
R: So what did you want to say?
A: when he meets-- when he faced! When he, come out, with some new vocabulary. Aiii, which one better?
R: mm?
A: when he, when he met. When he meets. When he, what did I say? When he meets some, new vocabulary. (listens to recording), ya meet. Look it up.

PT 3 If he come across some new word

SR 3
A: he, he comes across, or he come? (listens again) If he come across. If, so no “s.” No? If he will, if he comes across.

Aurora’s Sessions followed a pattern that was the same throughout from Session 1 to Session 3. Overall, she claimed that her confidence in speaking had increased, and she thought
she had become “much better” in her presentations because she now had strategies for note-taking, rehearsal, and time management.

**Grammar: Candy and Ella**

Candy and Ella were in the same class as Aurora. Like Aurora, they identified many Grammar items in their Sessions. While Grammar constituted 68% of all the items Aurora identified in her Sessions, Ella and Candy identified a smaller proportion of Grammar items, 61% and 51% respectively.

From the background questionnaire and interview it is clear that Candy believed that her English only started developing at Form 2, although she had been studying English since primary school. She preferred using Cantonese because she could easily express herself in it. Most of her daily interactions were conducted in Cantonese. As well, e-mails to friends and relatives were written in Chinese. Overall, Candy did not have much confidence in her English language abilities. She thought she did not have enough vocabulary to write well, and reported that she had not had many positive experiences using English: “I don’t have enough vocabulary, and my, grammar or my tense is, like a mess so I don’t like to, talk, with others in English.”

Though she did not mind speaking English in school, she disliked that it was imposed by teachers in some classes, and resented being made to stand because she had broken the “English only” rule in a lesson. She felt that there were times when using Cantonese to clarify something would have been more convenient and efficient. Together with Aurora and some other classmates, she instituted “English Days.” Once a month, the group would only speak English the entire day, even when they went out to tea after school. She found it a lot of fun, “because

---

16 Form 2 is the second year of secondary school, when students are between 13 and 14 years old.
17 A distinction is made between Cantonese and Chinese. The spoken form of the Chinese language used by the participants is Cantonese, while the written form is referred to as Chinese, which is standard among Chinese users, regardless of regional varieties of spoken language.
Sometimes when we are facing some difficult words we are just, using the Chinglish,\textsuperscript{18} and then we will laugh.”

Similarly, Ella preferred using Chinese because she could express herself better in it. However, she wrote most of her e-mails in English because she could type faster in English. She liked speaking English in school because she seldom had the chance to do so outside of school. She recounted how delighted she was when she discovered she could communicate with Japanese students using English while touring Japan with the school choir. Unfortunately, she had had some negative experiences using English outside school in Hong Kong. Some of her peers commented that she was trying to show-off.\textsuperscript{19} She liked the English oral exam because it did not require studying. Nonetheless, she felt that she was better at writing than speaking in English because when she spoke, she gave little consideration to grammar and tense. If she did, it would take too long and “people would have moved on to a different topic.”\textsuperscript{20}

Both Candy and Ella found participation in the study useful and fun. Through transcribing each other’s presentations, they discovered mistakes they had in common, and others that they had not been aware of. Candy said: “I found out all my mistakes when I am transcribing my speech. I realized that I have got many grammatic when I’m delivering my speech. My listening skills had also improved.” In addition, Ella discovered that there were some mistakes she made fairly consistently. From transcribing herself, she realized that she added “-s” randomly to the ends of words. Now that she had become more aware, she tried to reduce the number of times she made this mistake.

In the Pair, Candy was more lively and talkative while Ella tended to be more reticent. Nevertheless, over the course of the Sessions, a relationship developed such that Ella was

\textsuperscript{18}“Chinglish” is a compound word formed from “Chinese-English” used in Hong Kong to denote a heavily Chinese influenced form of English, which is generally ungrammatical.
\textsuperscript{19}Ella gave an example in Cantonese as follows: 奥，你扮嘢啊? 而家好叻啊你式英文, which translates to “Oh, what are you pretending to be? So you think you’re very smart because you know English?”
\textsuperscript{20}The comment was made in Cantonese (唧係我諗得來人地都已經講完嗰個話題).
contribute equally to the discussion by the end. Both spoke positively about working together after the Sessions were completed. Candy said that “if we encounter any problems, we can discuss with each other immediately.” She also mentioned that she liked having Ella there to exchange ideas with because “she provide some vocabularies for me and it is quite good.” Similarly, Ella saw Candy’s presence as an advantage because “my partner can point out my mistakes immediately and suggest how I can present my speech in a better way.” She noted that Candy was forthright and would not hesitate to tell her about her mistakes. Initially, Ella found it awkward to receive such direct feedback. However, she came to a matter-of-fact recognition that a mistake simply required correction, and there was nothing more to it. After that, she felt more positive about the process.21 After her first Session, where there was little interaction between the participants, Ella realized there was little point for Candy to be beside her if they did not interact acknowledging that interaction was learning.

One of Ella and Candy’s strategies for solving problems was to use the original prompt passage. As the participants become familiar with how the Sessions were conducted, they took advantage of the opportunity they had to consult the prompt passage after SR 1, as shown in the example below. They adopted this strategy at Candy’s Sessions 2 and 3, as well as Ella’s Session 3.

Example 7. (Ella, 56, Session 3, SR 3, S-agree)

E: “She neither spends time, on examination drills, nor,” (both laugh) I don’t xxx “attends,” “nor attends,” “nor attends?” “Nor attending any…”
C: nor does she,
E: “attend,” “attends”
C: uh nor she
E: “nor attending any”
C: or “nor attending”
E: which one?

21 Ella said in Cantonese: 啊唔係喔，係 唔真係錯呀嘛，嘅講出來唔係 ok 嘅覺得，改就ok 唸。咁嘅啦，我覺得。(Translation: oh but indeed, there truly was a mistake. So speaking out about it is ok I feel, and it just requires correction. It’s ok now, quite good, I feel)
C: “she didn’t sp-,” well “she neither spend-s time on examination nor attending, nor”
E: if we change it in a,
C: (laughs)
E: but I thinks, “attending” is… I donno. “Nor spend”
C: you see, (She points to the prompt passage) “she does not spend time on examination
drills like many nor does she attend, nor does she attend, tutorial classes”
E: ah! That means, I don’t need to use the “neither.”

There were instances, however, when working together was frustrating. For example, when each had some idea of what the correct form should be, but neither was certain, verbalizing did not lead to a resolution. In addition, Candy said it was also not challenging at times because they kept discovering the same mistakes over and over. The only difference was that they were noticing and correcting them faster. This helps us to understand the quantitative data, which show an increasing number of items being noticed as Type 2 Engagement, and an increasing number of JP Trackable items that were FN(Correct) across the Sessions.

As noted earlier, while each of Aurora’s Sessions followed the same pattern, the interaction between Candy and Ella changed over time. Table 32 below shows the differences between Aurora, Candy and Ella during the preparation time in each Session. Preparation time was when the participants could choose what they wanted to do. For Pairs, they could talk with each other, or remain silent, as the participant who was presenting focused on getting ready to present. For Individuals, as they were alone, the participants would be rehearsing the passage quietly.
### Table 32
**A Comparison of What was Said During Preparation Time Before PT 2 and PT 3**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Session 1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before PT 2</td>
<td>Before PT 3</td>
<td>Before PT 2</td>
<td>Before PT 3</td>
</tr>
<tr>
<td>Aurora</td>
<td>- Nothing said</td>
<td>- Nothing said</td>
<td>- Nothing said</td>
<td>- Nothing said</td>
</tr>
<tr>
<td>Candy</td>
<td>- Nothing said</td>
<td>- Candy solicited Ella’s help with adding information</td>
<td>- Nothing said</td>
<td>- Nothing said</td>
</tr>
<tr>
<td>Ella</td>
<td>- Nothing said</td>
<td>- Candy solicited Ella’s help with adding information</td>
<td>- Nothing said</td>
<td>- Nothing said</td>
</tr>
</tbody>
</table>

**Session 2**

<table>
<thead>
<tr>
<th>Before PT 2</th>
<th>Before PT 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Nothing said</td>
<td>- noticed what was written in the original prompt</td>
<td>- Nothing said</td>
</tr>
<tr>
<td>- Nothing said</td>
<td>- noticed an item</td>
<td>- discussed sentence structure and preposition</td>
</tr>
</tbody>
</table>

**Session 3**

<table>
<thead>
<tr>
<th>Before PT 2</th>
<th>Before PT 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Nothing said</td>
<td>- Ella noted correct form in the prompt passage. Candy asked for help with pronunciation.</td>
<td>- Candy made a suggestion; Ella asked for help for about preposition.</td>
</tr>
<tr>
<td>- Nothing said</td>
<td>- Ella supplied correct tense to Candy’s question. Candy asked about preposition.</td>
<td>- Talked about being fluent. Candy suggested she should rehearse more.</td>
</tr>
</tbody>
</table>

I found the increased interaction between Ella and Candy notable. At Ella’s first Session, Candy spoke more even though Ella was the focus participant. Below is an excerpt from Ella’s first Session in which Ella’s reticence can be seen in the long pauses.

**Example 8. (Ella, 5, Session 1, SR 1, verb)**

C: ae Ella have changed live for about a year to “have already live-ded”
R: mm-hm. Have already lived. Have already -
C: lived for about a year
R: have already lived for about a year. ok. And, did you make a change?
C: no.
R: ok. Ella, tell me about your change. Why did you make the change?
E: Uh because, em (6.0) um I donno I just think there is some problem and I wanna make changes.
R: mm-hm. So why is the change better than the original?
E: mm (6.0) mm (13.0) because the-s still living in Hong Kong and, and (3.0) mm (5.0)

However, things began to change over time as the Pair started to develop a joint focus on the introductions and conclusions of their presentations, and on how to become more fluent.
They became more invested in each other’s work, and Ella started to speak up. Beginning from Session 2, I observed Ella take the initiative by asking for Candy’s help, or drawing Candy’s attention to something.

Example 9. (Ella, 35, Session 2, SR 2, meaning/lexis)
E: do you feel it is strange in the information and technolog-
C: yes ah!
E: technology oriented age (laughs)
C: (laughs)
E: but I don’t know how to change it. Uh. And uh, and the, electronic age
C: uh may- yeah yeah
E: we change it
C: yes
E: in the electronic
C: age
E: how is it?
C: it is better than “in this information and technology”?

Over time, the Pair also recognized the mistakes that they frequently made. In Ella’s final Session, Candy pointed out that Ella had once again made a mistake with a noun, adding an ‘-s’ where it did not belong. The example below illustrates this.

Example 10. (Ella, 52, Session 3, SR 1, noun)
Ella’s transcript
Joyce Chow is a form seven students (student).
C: and here, students
E: uh again
C: again ya
R: what do you mean again?
E: the “s”
C: “s” again (laughs)

I observed that Candy and Ella’s interactions mediated their understanding of the rule for modals. In Ella’s Session 2, at SR 3, Candy pointed out that Ella had made a mistake and said “And, the-, websites can remembers which web pages, did you visit, by leaving a small file called cookies.” When I asked what the form of the verb “remember” should be, they replied in
unison “can remember!” In the following Session, Ella’s Session 3, each time Ella added an ‘s’ after the modal, Candy was quick to mention it.

Example 11. (Ella, 52, Session 3, SR 1, modal)
Ella’s Transcript And she can also learn (learns) new words
C: here. She can also learns, you change it, this learn to learns
E: mm. He-ha!
R: but what should it be?
E: she can, learn
C: learn. Is it there is no “s” after “can”?
E: uh-huh I, I donno
C: (laughs) … You, you choose
E: um, I change

Following this, Candy continued to point out when Ella added “s” to the verb after a modal. Here are some examples:

Example 12. (Ella, 64, Session 3, SR 1, modal)
Ella’s Transcript she will learns all the words
Candy’s Transcript she’ll learns (learn) all the words.
C: she will look up the dictionaries and she will learn, learn
E: uhn
C: because you have delete “she will” so is
E: learn

Example 13. (Ella, 75, Session 3, SR 1, modal)
Ella’s Transcript And she thinks that these chatting on the net can also trains her typing skills.
Candy’s Transcript And she thinks that this chatting on the net can also trains (train) her typing skills.
C: aw you, change the word to avoid, aw here
E: chan- change
R: so
C: can, because of the can
E: can train
For the rest of the Session, Candy continued to monitor Ella for this frequently made mistake. She, for example, commented that “I still heard some ‘s’ after the can, after can,” and tried to help by reminding Ella before PT 3.

C: do you remember that it did that, after can there’s no “s” uh? Are you sure?  
E: (laughs) I’m not very sure, but I think it is  
(both participants laugh)

At PT 3, Ella succeeded in correcting several of the modal forms that she discussed with Candy. Though some of the forms were not re-used, verbalizing the mistakes Ella had made had the effect of solidifying the rule in Candy’s mind, who in turn, mediated Ella’s understanding of it.

_Pronunciation extended example: Bobo, TC, and Swatch_

Bobo was a F.7 student who was assigned to the Individual condition. She identified the highest total number of Pronunciation items in her Sessions and was the only participant for whom Pronunciation items constituted the largest proportion (31.2%) of all the items she noticed. She had a strong sense of her identity as a Hong Kong Chinese whose mother tongue was Cantonese. On the whole, Bobo enjoyed speaking English in school and took advantage of the many opportunities there were to do so. Like many of the other participants, she reported thinking about schoolwork in English if the subject was taught in English, but in discussion of school projects with friends, her preference was to use Cantonese, except for key academic concepts and terms that she had learned in English. Similarly, since the school was a missionary school, and she had learned prayers and hymns in English, she preferred to use English for religious worship.

In the exit interview, Bobo spoke extensively about what she had learned from transcribing herself. Using Cantonese, she told me she had never listened to herself speak
English before the study, and that her first attempt to transcribe her own presentation was a great revelation. She had expected to be able to transcribe as she listened, but she discovered that the first sentence she spoke was error-filled and hesitant. She did not like how she sounded. It was an unexpected and unsettling discovery, one that she found hard to accept. Over the Sessions, she thought her performance fluctuated a great deal such that on some occasions, she was coherent, while on other occasions, her presentation was incomprehensible even to herself.

Bobo was critical of the English spoken by people in Hong Kong, and observed that “fundamentally, Hongkongers speak English strangely.” 22 She said that their speech was full of speech fillers such as “ae” and “ah,” and they typically failed to pronounce the final consonant on verbs such as “used to.” Since she identified strongly as a Hongkonger, her critical comments about Hong Kong English speakers suggested she rationalized the way she spoke. This was supported by her comment that she had not gleaned any new insights from transcribing others in the study.

Nonetheless, Bobo reported that she felt more confident about the oral exam after the study, even while maintaining that people did not really understand her because she was not able to fully express herself in English. If given a choice, she would have preferred to have collaborated with a peer-partner. She believed that the interaction would have been enjoyable and beneficial.

What Bobo noticed did not change much between Sessions. The items remained superficial differences (DIT items), or disagreement over how she had been transcribed. In Session 1, she noted at SR 1 differences in where she had marked pauses in her transcript in comparison to the other transcriber. At Session 2, she shifted her focus from noting DITs to doubling the proportion of Pronunciation items noticed. Of these, more than half were Deny

---

22 This is what Bobo said in Cantonese: “我會覺得始終香港人講英文真係好奇怪。”
items, where she disagreed with how the other transcriber had rendered her pronunciation. At
Session 3, she continued to identify a high number of Pronunciation items. An example is shown
below. I did not notice any mispronunciation at PT 2 or PT 3.

Example 14. (Bobo, 44, Session 2, SR 1, deny)
Other’s transcript the recycling bins in the next “willy”, “willage” (village)

SR 1
B: will, will-age. Willage. … she, thought that my pronunciation is wrong.…. I said Willage
… not village… which is a very-
R: so which one is correct?
B: I heard myself (laughs) said “village”, not “vill.”

PT 2 The family will carry large bags of wa- of bottles and cans to the recycling beans you
know next village,

PT 3 the recycling bins in the n-next village,

Many of the items that Bobo identified were not re-used at PT 2 or PT 3, and of those that
were, Bobo correctly pronounced most of them. However, of those with which she had difficulty,
verbalizing the item did little to mediate her production without correct input from another
source. The difficulties she had with pronunciation could have been clarified had she been
working with a peer-partner whose transcript she would have been responding to. The example
below illustrates this.

Example 15. (Bobo, 43, Session 2, SR 1, deny)
Other’s transcript this is because of her “mara’s” (mother’s) influence,

SR 1
B: this word because of her, I said maather’s (mother’s) influence
R: mm-hm
B: (laughs) How come she? Maather’s influ-.

PT 2 Angela said that this is because of her maather’s influence,

PT 3 her interest in protecting the environment is due to her mother maather influence,
**Pronunciation: TC and Swatch**

In contrast, TC and Swatch worked as a team. Together, they were more effective than Bobo in their approach to pronunciation items. TC and Swatch were classmates, but had not been close friends prior to the study. Both of them thought the exam practices increased their confidence. TC said that she used to be nervous when she spoke, but now she felt more confident. For Swatch, who got a low grade in the last oral exam, the English speaking exam was something she did not like working on. However, she reported, she had since changed her mind and was looking forward to getting a good grade in the upcoming exams.

TC remarked that transcribing herself made her aware of her grammar and pronunciation mistakes as well as the great speed at which she spoke. She had not realized that her pronunciation was problematic until she transcribed herself, because she had always been able to communicate with others. From listening to herself, she grasped that it would be difficult for her listeners to understand her rapid speech. From transcribing Swatch, TC noticed how Swatch created introductions and conclusion in her presentations. She also discovered that Swatch was adding ‘s’ to words haphazardly, a problem they had in common.

In Swatch’s case, she discovered through transcribing herself that she tended to repeat the entire sentence when she made a mistake, which drew attention to the mistake. From transcribing TC, Swatch became aware that in order to sound more interesting, she needed to work on her intonation and rhythm. When asked what they had changed in their speaking from participating in the study, both TC and Swatch wrote that they learned that they should speak slower. TC said that she had to pay attention to when it was correct to add “s” to a word, and vary the use of the word “that” in her presentations. Swatch said she had become more aware of tenses. When they were asked what they valued most in working with a partner, both noted that the other could help
them notice their mistakes in pronunciation and grammar. The examples below illustrate how TC mediated Swatch’s learning of the pronunciation of “contact.”

**Example 16.** *(Swatch, 9, Session 1, SR 1, accept)*

TC’s transcript  My topic “Fear of content.”

**SR 1**

T: ya I hear contact content instead of contact…
S: this is correct? Are you sure?
T: no, no I hear this one
S: hah? There is different
T: content, contact
S: con- content
T: your, the correct right should be contact right
S: yes
T: but I hear content
S: aw.

**PT 2** My topic, is fear of content

**SR 2**

S: I’m afraid of content, afraid of, con- afraid of content …
T: contact contact
S: afraid of content, contact
T: I think this time you say it correctly
S: huh, great problem with the word.

**PT 3** My topic, is, fear of contact

Another way in which peer-partners supported the participant under focus was to remind her of errors that she frequently committed, but did not notice. In the example below, also generated at SR 2, we find TC reminding Swatch of a recurring error in pronunciation. The outcome of this exchange was positive. Swatch used the word “foreigner” five times at PT 3, and they were all correctly pronounced.

**Example 17.** *(Swatch 11, Session 1, SR 2, accept)*

T: forjiners uh (laughs)
R: say that again?
T: she say for-jiners
R: again?
T: yeah
R: did you notice that?
For TC and Swatch, speaking slower for clarity was a common thread running through all of TC’s presentations. At Swatch’s last presentation, I observed TC advising Swatch to slow down for greater precision in her pronunciation. It appeared that TC had internalized the importance of speaking slowly, and was now reminding Swatch instead of the other way around.

In contrast to Bobo, neither Swatch nor TC identified many DIT items, and over the Sessions, the number of Pronunciation items identified by either of them decreased. Like Candy and Ella, over the course of the Sessions, TC and Swatch developed a relationship that enabled them to trust and respect each other’s contribution. In the exit interview, when they were asked separately what they would target specifically to follow up, both mentioned the pronunciation of the “-ed” and “-d’ word endings. In addition, Swatch noted the “s” sound that seemed to appear at random in her speech.

In sum, engagement and its effects differed between Individuals and Pairs. Individuals proceeded along the same path at each Session. Aurora remained focused on grammar, while Bobo kept noticing DIT and pronunciation items. While it was helpful to talk, and Aurora found it a useful strategy to decide what change to make based on whether it sounded correct, there was little that Bobo could do without input from other sources to correct her pronunciation.

It can be seen from the extended examples that participants in Pairs were able to take advantage of a peer-partner who shared resources and offered feedback. The peer-partner, through talking, mediated the difficulty encountered by the participant, and at times, gave helpful reminders about frequently made errors. The interaction between the members of a Pair was
mutually reinforcing. The members of the Pairs also developed together, through the Sessions. In addition to sentence level editing in the later Sessions, they began to focus on other aspects of the speaking presentation, such as the structure of the speech and strategies for fluent performance.

Section Two

In this section, I highlight some other differences between Individual and Pair interaction that are not illustrated in the extended examples.

First, there was a great deal more variation in what was said in the Pairs. In length, they ranged from short utterances like those of Individuals, to lengthy, reiterative, and protracted conversations over a single grammar point or idea that the Pair was working to resolve. The presence of the peer and the relationship between the two was the main cause for the variation in the verbalizations. This was because the peer-partner played many roles that supported the participant, which critically shaped the Sessions. Sometimes, the peer-partner acted as a team mate, helping to solve a problem or co-construct new knowledge. At other times, she played the role of the more knowledgeable tutor leading the participant to a correct solution for a problem. Often, she helped to draw attention to something which led to closer examination and review. Peer-partners were also cheerleaders who offered encouragement, and whose opinion and affirmation were valued when the participant under focus was uncertain about something. Individual participants did not have such a resource.

As mentioned, peer-partners also made a difference to the timing of an item being noticed. I report in Chapter 4 that their presence accounted for a higher number of items noticed being at SR 2. In the following example, Congee is a source of information for vocabulary, helping Agatani express her meaning. We can see the improvement at PT 3 following this noticing episode.
Example 18. *(Agatani, 40, Session 2, SR 2, S-meaning/lexis)*

A: … you’ll been captured, in the front door …
C: catch, in the entrance
A: entrance, but entrance is in- is it?
C: try lock (unintelligible)
A: if you- we miss you go out is exit. But entrance is like 入口, exit is 出 23
C: lobby, lobby of the--
A: uh lobby, good (laughs)

**PT 2** When you exit through the front door, there is a, camera taking your video also.

**PT 3** And when you arrive the lobby, you will find one or two camera pointing to you.

One Pair in the study was notable for rehearsing the entire passage before attempting PT 2 and PT 3 recordings. Froggie and Dull used each other as “live” sounding boards offering immediate feedback and reminding each other of the items that should be corrected. As they rehearsed, they were able to examine and revise the presentation, which created an additional opportunity for noticing problems in sentence phrasing and gaps in content. In the example below from Dull’s Session 3, she solicits Froggie’s assistance in developing a concluding opinion for the presentation. Dull identified this need at SR 1, but it was in returning to it over and over between working on other items during the Session that a solution was finally found. The full discussions from each Session are presented in Appendix L. The improvements and refinements made at each stage can be seen in the presentation excerpts below:

**Dull’s transcript**

And in my opinion, I agree with his point of view, I think that learning English is in, we cannot improve our English by only spending some days but we have to practice it all throughout. And then read and write is (are) very important so thank you.

**PT 2** In my opinion, I think, English can’t be improved, within, short, period of time. If we want to make a proo- improvement, we have to, read- more- and- write- more.

**PT 3** In my opinion, I think that English can’t be improved, within a short period of time. Instead of acting pacsif- passively, we should, take the initiative, to-- read- more- and- write- more like, Victor, Victor Lee does, every day.

---

23 入口 is ‘Entrance’ in Chinese. 出 is an incomplete utterance for ‘out’ or ‘exit’.

124
Over time, I noticed that the members of Pairs became familiar not only with the format of the Sessions, but also with their peer-partner. The peer-partners developed a camaraderie that led to inside jokes, reminders about each other’s habits, and recurring items that each in the Pair needed to pay attention to. In fact, awareness of some of their partner’s difficulties extended beyond the participant’s Sessions. Daydreamer was so aware of her partner’s inclination to use “he” regardless of gender that she burst out laughing when she heard Kitkat making the mistake in her own final Session.

At times, it was difficult to know what the participants were thinking when they said to each other “you know what I mean?” Being classmates sharing a learning environment beyond the Sessions afforded opportunities to find common ground. For instance, peer-partners would use stock phrases or key terms from other subjects that they studied, such as “spread the news” from Bible study or “biodegradable” from biology, and because these were used in a novel context, the students shared a good laugh. There was no opportunity for this type of engagement during the Individual Sessions.

As seen in the extended examples, Pair participants viewed the opportunity to interact with a peer as highly positive. Thinking of solutions together was fun for Agatani, who thought that it brought her and her peer-partner closer as classmates. Participants also claimed that it “feels better to have someone to share” and “we can learn from each other’s” mistakes. Daydreamer thought that having a partner present made her less afraid and nervous. Amo noted, in Cantonese,²⁴ that having a partner meant that there was a chance for discussion, acknowledging that verbalizing between peers open up learning possibilities.

Some of the time-pressed participants said the worst thing about working in a Pair was finding time to meet. In spite of this, almost all (13/16) of the Pair participants said they would

---

²⁴ What she said was “起碼有得傾.” The literal translation is “at least, there is discussion.” Figuratively in Cantonese “有得傾” (discussion) could mean negotiation or possibility.
have come to different conclusions and solutions if they had worked alone. Ella thought she
could have completed the tasks with greater efficiency and focus, but noted that she would not
have been aware of as many problems. Unlike Ella, Amo thought she would have taken a longer
time to complete the tasks, but would also have made fewer changes without a partner. While
most of the Pair participants were comfortable verbalizing with a partner, they admitted that it
took time to get used to the process and the presence of another person scrutinizing their work.
But once they were used to it, they found it helpful.

However, not all participants in Pairs developed friendships that opened up opportunities
for discussion. For one Pair, which comprised a strong student and a weaker one, while the
Sessions were cordial, it appeared to me that the stronger student was not as invested in the
process as the other. In the exit interview, the stronger student admitted that she had not been as
forthcoming as she might have been in pointing out all her partner’s weaknesses, and she thought
that the weaker student acted similarly. If she noticed something that she thought was not a
serious mistake, she would not note it specifically, but gloss over it by saying “it’s not very
fluent.” She explained she did not want to add to her partner’s nervousness and lack of
confidence by highlighting every mistake she made, as it could be considered mean spirited.25
This was also her justification for not suggesting ways of improving her partner’s performance.
Nonetheless, she thought that they developed a harmonious working relationship. The weaker
partner was more generous in her assessment. In the exit interview, she said she had gained a lot
more confidence in her presentation ability, and the presence of her partner was helpful because
they could solve problems together. Across Pairs there were reported advantages to pair work, in
spite of differences between members of Pairs in how the partnership was viewed.

25 She said:
“佢係好似好衰啲啲如果你係噉，佢係話佢唔度又係錯嘅度又係錯嘅話。佢係好似好衰啲啲。．．．驚佢會唔開心啲” (It’s
like it’s very mean if you, if you say it’s wrong here and it’s also wrong here. It’s like really mean. I’m afraid she won’t be
happy.)
In contrast, Individual participants did not have the advantage of working with a partner and sharing resources. Individual participants were aware of this, and almost all of them indicated at the exit interviews that they would have preferred to have had a partner. Some participants noted that it was difficult to work alone because it felt strange to talk to no one in particular. Bobo said she felt like she was trying to convince me that she was right during the stimulated recalls. On the other hand, Buzz said she preferred working alone because it simulated real exam conditions, and Zax thought the relaxed and more informal atmosphere that might prevail if a partner were present would have been a distraction. None of the Individuals considered their interactions with me during the Sessions to be a benefit or advantage. While I did at times ask for clarification or repetition when the Individuals were unclear or inaudible during the stimulated recalls, these were not intended to be prompts.

To summarize, there were many differences between Individual and Pair engagement. The main difference between Individuals and Pairs was the peer-partner support that was afforded participants in Pairs. Partners tended to develop camaraderie and friendship, which created a collegial and encouraging atmosphere to work within while remaining focused and on-task. The Pair partner became an important resource who shared their partner’s successes and failures. Sometimes, the teasing, or the prompt “Remember?” helped participants to keep things in mind with which they frequently had difficulty.

Most notable was the development of shared values and ideals about what the presentations should consist of among Pairs. They developed a working rhythm and used strategies that were reinforced over the course of the Sessions. Over time, threads of the peer-partners’ shared history from prior Sessions were woven into the current Session. Although each participant had her specific areas of focus, partners were often asked to help construct or to comment on discoursal features or devices to draw listeners in, or a clearer and more refined
conclusion and personal opinion. Fluency was a goal for many participants as well. These goals eluded quantification in the dataset because what were enumerated were individual instances of noticing that did not reflect the context, or the entirety of the presentation. I call these goals “macro-preoccupations,” which arose as a result of the joint focus shared by Pairs.

Section Three

Macro-preoccupations

While Individuals’ approached their Sessions in the same manner each time, Pairs evolved over the Sessions. Pairs tended to focus on areas beyond Grammar, Pronunciation, and Content as the Sessions progressed, and there was a tendency for their focus to converge on particular aspects of their presentations. These aspects were fluency, discourse elements, content, and exam strategy. Some partners adopted the same strategies and reminded each other of them. Some of the strategies that they adopted were rehearsing the entire presentation with the partner, not making major changes to the presentation, and keeping an eye on the clock. I call these evolving foci macro-preoccupations because they were frequently mentioned, and the participants judged their performances based on how they did in these areas.

Fluency and speech rate

Both Individuals and Pairs frequently pointed out where dysfluencies occurred in their transcripts and productions tasks. Fillers such as “em” or “ae” were considered ugly. Many participants recognized that these fillers were often involuntarily spoken, and they understood that they were filling the silence because they were trying to remember what to say next. Some participants considered pauses or “dead air” a sign of their lack of control, and as missed opportunities for the participant to present more content. The meaning of fluency changed for
some Pairs. Many Pairs started the study equating fluency with a high rate of speech. They soon realized that speaking too quickly led to stumbling, poor articulation, and incomprehensibility. Towards the end of the study, some participants judged their fluency by how “smooth” they sounded instead, which took into account dysfluencies and fillers.

Among Individuals, fluency was mentioned in 12 of the 24 Sessions by 6 out of 8 participants. For Pairs, however, this was mentioned far more frequently: in 36 out of the 44 Sessions and by every Pair participant. Also, only 3 out of 8 Individuals mentioned fluency at the last Session (Session 3), compared to 13 out of 16 Pair participants. In other words, there were more comments on fluency at the final Pair Sessions compared to the earlier Sessions, made by both the Pair participant and her partner. This was observed as Pairs suggested to each other that she slow down, or commented on each other’s effort or success at doing so.

Looking at the Pair transcripts chronologically, I could trace the emergence of a participant’s concern and see it recur in successive Sessions. For instance, Congee noted at Agatani’s Session 2 that she was speaking too quickly, which Agatani acknowledged. In her Session 3, Congee noted of herself that she had spoken faster than before, but Agatani considered the pace acceptable. Later in that Session, the Pair discussed speed again, and Congee explained that “we will keep on accelerating” when nervous. In Agatani’s Session 3, speed was mentioned repeatedly throughout. It was noted by Congee, which Agatani acknowledged, and then evaluated after PT 2 and PT 3, with the result that Agatani concluded that she needed to slow down.

Potato and Cathy’s their macro-preoccupation with fluency was tied to rehearsal and preparation. At Cathy’s Session 1, Potato advised her to think about the next sentence before she spoke so that she would use fewer fillers and avoid long pauses. This was a problem she herself faced, and Cathy mentioned this in the next Session. At Session 2, Cathy noted that her
performance was improving as she was becoming more confident, calm, and fluent because she had rehearsed and read out the entire presentation as part of her preparation. At the next Session, Cathy advised Potato to use this same strategy and rehearse with her, but this offer was not taken up. Potato concluded at Session 2 that she had become more fluent because she was “thinking better and linking sentences,” which reduced the pauses. She thought she could have added more content using the time saved. At Session 3, Cathy claimed that her fluency had improved because “I think then I speak,” which echoed Potato’s advice to her at Session 1. But now, in her final Session, she recognized that she lacked rhythm and flow in her speech, and she had a staccato pronunciation. Potato concurred and suggested that Cathy should try to speak in a “lighter, smoother way.” Thus, Cathy’s goals shifted over the Sessions beyond saying sentences without pausing. At Potato’s Session 3, she commented that her own performance had become more fluent because she was not thinking too much about the grammar. This directly contradicted her advice to Cathy at Session 1 to think about the sentence before she spoke. But it also indicated her awareness that thinking too hard about grammar inhibits fluency. Meanwhile, Cathy reinforced Potato’s claim that she had heard it said that examiners do not really pay attention to what one presented, but how fluently one spoke. It appears that the goal of sounding fluent when presenting took precedence over grammatical accuracy.

Fluency as a macro-preoccupation was also in evidence in the interactions of other Pairs. Kitkat and Daydreamer were concerned with the many fillers (ae, ah) that were found in their presentations. Their solution was to slow down. Froggie and Dull were similarly preoccupied with fluency. Froggie was cognizant of the break-neck speed with which she spoke. Before PT 3 in Session 1, in an almost conspiratorial whisper, she asked Dull “do you think I can, actually, present slower?” This remained a focus in Session 2 as she continued to work to slow down. At Session 3, she happily commented that she was “quite fluent” and “I intentionally, slower, speak
slower.” This was facilitated by Dull making hand motions during her presentation to signal her to slow down. In this way, Dull actively helped Froggie perform in her final Session.

**Organization of presentations**

Another macro-preoccupation was content and discourse elements. Many of the Pairs began in the early Sessions by mainly with noticing grammar items. Before long, however, they realized that good introductions and conclusions framed their presentations better, and an opinion offered at the end helped to personalize their speech. While these elements might also have been improved in the Individuals’ presentations, they were seldom mentioned in the Sessions, and I noted minimal change between the initial presentation and the final product.

In the Sessions with Dull and Froggie, the two were constantly working together to formulate a sensible conclusion and an opinion. Particularly Dull, from Session 1, she collaborated with Froggie to test out expressions and formulate her sentences, and at Froggie’s invitation “do you want to practice with me?” they began rehearsing with each other to test out the presentation. It was amusing to observe the extent to which Dull saw Froggie as an integral co-producer of her presentation, and to hear Froggie reminding Dull about ownership when she asked “I dunno, what do you think?” Froggie replied, “It’s your speech, not my speech.” Over the next Sessions, the roles reversed. Dull invited Froggie to rehearse with her, and Froggie asked Dull to help her conclude her presentation “in a better way.” Forming coherent conclusions and including opinions became a macro-preoccupation for this pair.

Also focusing on introductions and conclusions were Congee and Agatani. From Agatani’s Session 1, it was clear that she was willing to take a lot of risks by not adhering closely to the prompt passage. She was also forthcoming with ideas for Congee’s presentation at Session 2. While Congee did not immediately adopt all her ideas, she adopted Agatani’s suggestions for a
more arresting introduction in Session 3, and even noted this with appreciation in her exit interview: she said her presentation had improved because of Agatani’s contribution.

**Exam strategy**

A last macro-preoccupation that was observed in the Pairs was strategy. As mentioned earlier, Candy and Ella told each other “don’t change so much,” and they made few structural changes over the course of their Sessions. They adopted this strategy because they found sentence revisions difficult to remember, and caused them to pause in their presentations when they tried to execute them. Rehearsal was a strategy shared by Dull and Froggie that worked to their advantage.

The F.7 Pairs who would soon face the public exam tended to concentrate on exam skills and techniques. For example, Swatch and TC mentioned “time control” in their Sessions, in relation to ensuring they could say all they had intended, including the articulation of a coherent conclusion in the allotted time. When TC ran out of time, Swatch commented that she needed better “time control,” and suggested, in different Sessions, that “if time is not enough, just don’t read out those points.”

I observed that Amo was looking at the timer throughout her initial presentation, instead of her audience, at Session 1. She stopped when this was pointed out to her. Amo had many insights into the exam, which she shared with Pygpyg. While Pygpyg attempted to include as much of the original passage in her presentation as she could, Amo suggested that content was not the most important object of the exam, skill was. Over the Sessions, Pygpyg followed this advice by shortening her presentations in both Session 2 and Session 3, and trying to formulate concluding opinions instead. As Pygpyg made the changes, Amo also advised her not to “change too much.’ This was clearly a strategy she followed because in her Session 3, there was
negligible difference between PT 2 and PT 3. She was working to refine her articulation and delivery instead of her content. One way to concentrate on the articulation of her speech was to minimize the paraphrasing of the passage. Particularly when she was not sure of its meaning, she would copy the exact wording of the passage. At Pygpyg’s Session 3, another exam strategy that Amo suggested was to not react if she made a mistake, as that would bring attention to it. Instead, Amo suggested to Pygpyg that she “pretend that you have said everything correctly.” I observed Amo doing this at her own Session 3: she smoothly carried on and did not draw attention to a mistake she had made. To help Pygpyg lower the risk of mispronunciation, Amo suggested that she use common words rather than try to impress the examiners with words that she was not sure how to pronounce.

Another strategy articulated by F.7 participants was in regard to content accuracy. Cathy, Potato, Swatch, and Amo thought that a little inaccuracy in the content was tolerable. For example, Swatch commented that “it doesn’t matter, the teacher won’t know.” For these students, the “how” of the presentation was more important than the “what.”

In sum, Pairs benefited from working together on macro-preoccupations in ways that Individuals could not. As the Sessions progressed, some Pairs developed a therapeutic dynamic that included a joint focus. The macro-preoccupations pervaded subsequent Sessions which allowed for continuous improvement in the participants’ presentations.

**Chapter Summary**

In this chapter, I discuss the differences in engagement between Individual and Pair participants through the use of extended examples, and highlight other differences and macro-preoccupations not illustrated by the extended examples. The differences between Individuals and Pairs can be seen in the quantitative analysis, while the qualitative analysis enriches our
understanding of the participants’ engagement and their effects. The presence of the peer-partner was facilitative and supportive in many ways. Some participants in the Individual condition indicated that they would have preferred to working with a partner.

Macro-preoccupations were difficult to glean from the quantitative analysis, but they can be traced through the Sessions in the qualitative analysis. These preoccupations reveal what was motivating the participants as they revised and re-attempted their speeches. At first glance at some of the final products, I was puzzled to note that nothing much had changed, particularly in the structure of the presentations. However, through an examination of the qualitative data, it can be seen that participants had deliberately decided to minimize changes in the order of the information or structure of the passage in order to focus on articulation and fluency.

As well, the qualitative data reveal that the participants were attempting to refine presentation aspects that were difficult to capture quantitatively. For instance, many participants reported that they had improved their performance because they had become more confident. However, confidence is difficult to quantify. Many of the participants also commented at the end of their Sessions that PT 3 was an improvement over the earlier attempts at the presentation. To verify their positive comments, and to find out if any noticeable improvement could be discerned, I enlisted the help of an experienced Hong Kong teacher\(^{26}\) to rate their performances.\(^{27}\) Her ratings were a close match to what the students considered to be their best attempt at the presentation, which was usually PT 3.

Thus, there were differences in how Individuals and Pairs verbalized. For Individuals, verbalizing appears to have had effects that were facilitative to the task within the Session. However, without a peer-partner, Individuals could not benefit from the resources that a peer-

\(^{26}\) This teacher has experience grading the HKAL exam.
\(^{27}\) Using a random number generator, I sampled 12 Sessions (17.6%). For each Session, I scrambled the order in which the presentations were listened to, and then asked the teacher to rate, impressionistically, which was best and which was worst. Eleven out of 12 of the Sessions (91.7%) were rated where PT 3 was the best performance or were an improvement over PT 2. Only in one case was PT 3 rated the worst performance.
partner would have offered, nor were they party to the kind of evolving joint focus observed among the Pairs. These macro-preoccupations shaped the later Sessions with Pairs, and directed their noticing beyond Grammar, Pronunciation, and Content in ways that had a meaningful and positive effect on the task and performance. In the next and final chapter, I conclude my discussion of the research questions.
In this final chapter, I begin by discussing the results of this study in light of the literature reviewed in Chapter 2 by research question. Although the design of this study is more complex and less classroom-oriented than the studies reviewed, some of the data generated are comparable. In the second section I suggest some pedagogical applications arising from this study, and in the third section I discuss the limitations of the study. I conclude with suggestions for future research.

**Section One: Theoretical Contributions**

*Approaching noticing through collaborative dialogue*

As noted in Chapter 2, this study draws on Schmidt’s (2001) definition of noticing, which is noticing instances of language without abstracting rules or principles (p. 5). How noticing actually works continues to be intensely researched, mainly through a cognitivist lens. In this regard, Schmidt’s comment in his (1990) paper is prescient. Then, he noted that researchers have tended not to focus on the learners and what they were experiencing and thinking, but focussed on the products and outputs that are attributed to their actions. As such, investigating noticing through collaborative dialogue longitudinally in this study is my attempt to address this gap.

In working together, my study has shown that Paired learners often begin with noticing surface structure, no differently than Individual learners. However, over time, this joint attention becomes transformed. Pairs jointly create a shared memory of frequent errors, and focus on common goals as they collaborate. Without considering noticing from the point of view of collaborative dialogue, it is difficult to discern how this act is transformed into sustained joint attention, from the initial act of perceiving something.
As well, this study contributes to the existing literature on collaborative dialogue. The data in my study accumulates data that exemplifies the role of language in mediating learning through collaborative dialogue. Like Storch (2007), this study shows that peers exert pressure on each other to perform better. As well, my study, like others before it, suggests that the quality of engagement is a factor in achievement and performance (Kim, (2008); Storch, (2007); Watanabe and Swain, 2007). As a longitudinal study, the results are consonant with Shehadeh (2011) and Yarrow and Topping (2001). The participants in this study showed improvement within each Session, kept each other on task and gained confidence as a result.

Section Two: Comparing Results

Research Question 1

Research Question 1 asked what learners notice of their spoken language production. As discussed in Chapter 4, the participants noticed a large number of items not only in language form, but also content. The majority of items noticed were Grammar items. This is comparable to the findings of Lynch (2001) and Stillwell et al. (2010). In both of those studies, the majority of the changes the students made from Transcript 1 to Transcript 2 were changes in grammar. The proportion of grammar items noticed were 35.7% and 47.8% for Lynch (2001) and Stillwell et al. (2010) respectively. In this study, the proportion of comparable Judgement Possible (JP) Grammar items noticed was 63.7%. While Lynch (2001) reported that his students corrected predominantly verb tenses and article choice, the participants in this study made the most changes in subject-verb agreement and count nouns, consonant with Stillwell et al.’s (2010) study.

The participants in this study also noticed a large number of items in other areas, such as Content and Pronunciation. The proportion of JP Content items that were noticed in this study
was 19.7%. This is consistent with Stillwell et al’s (2010) study, where the finding for comparable data was 19.6%. For Lynch (2001), however, the proportion was higher at 26.8%. For Pronunciation, the proportion of items noticed in this study was 16.6%, compared to 19.6% for Lynch (2001) and 24.3% for Stillwell et al. (2010). It should be noted that Lynch (2001) and Stillwell et al. (2010) reported a “Mixed” category, which comprised 17.8% and 8.3% of their data respectively. There is no corresponding category in this study, which could account for the differences found between the studies. Another reason for the differences could be that this study provided more opportunities for noticing and SR through its many stages.

**Research Question 2a**

Research Question 2a asks what learners do with grammar, pronunciation, and content items when they notice them. When an item was noticed, a participant could act on it in several ways, for example, repairing it correctly. In this study, from Chapter 4, noticed items were correctly repaired in most cases. The proportion of correctly repaired Grammar, Pronunciation and Content items were 85.8%, 69.0%, and 76.8% respectively. On average, 81.2% of noticed items were correctly repaired. In comparison, Lynch (2001) found that 72.3% of the changes were for the better. In the study conducted by Stillwell et al. (2010), the number of accurate corrections increased from Presentation 1 to Presentation 2. For Presentation 1, 55.1% of the items were changed for the better, while in Presentation 2, this figure increased to 60.1%. Stillwell et al. (2010) also found that students improved in their ability to identify and correctly repair the items. This result is similar to the finding of the present study that the participants increasingly repaired the noticed items correctly at First Noticing, as reported in Chapter 5.

Some changes that the students made in this study were simple (usually morphological, involving one word) and mechanical, for example the addition or deletion of an “-s” or “-es” to a
noun or an “-ed” to a verb to accurately reflect the plural form or mark the past tense without involving other parts of speech in the sentence. While rule application was sometimes used, it was not a consistent strategy because the participants’ grasp of the rules was often incomplete or incorrect. Some participants often relied on what “sounds correct” or what was more intuitively appealing, reminiscent of the adolescent learners in Swain and Lapkin (1995, 1998). Like the adolescent learners in that study, most of the participants in this study experienced language learning as a communicative rather than a rule-based activity.

    The participants also tried to solve language problems using cognitive processes similar to those reported in Mennim (2012). Like Mennim’s participants, who relied on dictionaries and their peers as sources of information, the students in this study used the original prompt passage as a source of correct usage, and their peers as sources of information to help with their corrections. While Mennim noted that his participants drew on their L1 to confirm message meaning, this strategy was rarely used among the participants in this study. They did, however, make cognitive comparisons between problem forms and their existing knowledge of English, as Mennim’s participants did.

**Research Question 2b**

This research question asks what learners do with the items they notice when given an opportunity to repeat their speech performance. From Chapter 4 (Table 10), 65.5% of items that had been correctly repaired at First Noticing were correctly produced at the next opportunity. Compared to Content and Pronunciation items, more Grammar items were incorrectly produced after they had been accurately corrected. In Lynch (2001) and Mennim (2007), what was noticed was not compared with what was subsequently produced. Lynch (2007) found that participants in the Student-Initiated (SI) group succeeded in re-using the highlighted forms correctly at a rate of
64%, while students in the Teacher-Initiated (TI) group re-used the highlighted forms at the lower rate of 47%. It should be noted that the SI group followed a procedure that was similar to that used in this study in that the students chose what they wanted to focus on. In Mennim (2003), the study reported uptake of the teacher’s repairs in article use, preposition use, and pronunciation only. The number of items that were correctly revised and correctly produced at the final presentation was not reported. In contrast with this study, Stillwell et al. (2010) noted that many of the students did not incorporate the corrections in their second presentation. Of those that did, the students were more likely to use the teacher’s corrections (57%) than their own. About 55% of the corrections supplied by either the students themselves or the teachers were accurate. This is slightly lower than the proportion found in this study.

At this juncture, it is should be noted that though a large proportion of items were noticed, corrected, and correctly produced at later opportunity, there was still a sizable proportion of items that were not correctly produced (15.8%). Thus, the data show that noticing alone does not ensure a positive remediated outcome in the future. This is because it is just one step, albeit an important one, in the cycle. The choices made after noticing occurs are many, and depends on a myriad of factors ranging from memory, confidence, time of day, oral articulation, to personal choice and agency. Even then, it is not always clear what priorities or personal decisions have come to affect subsequent attempts. There is no straightforward one-to-one correspondence, nor is there a predictable formula for determining how a noticed item will be produced subsequently.

**Research Question 2c**

Research question 2c asks whether items that have been noticed are produced more accurately over time. As noted in Chapter 5, most of the items that were corrected remained
accurate over the course of the Session. However, there was insufficient data to support extrapolations to longer time frames.

In contrast, Lynch (2007) found that for the SI group, which selected a portion of the presentation to transcribe themselves, 64% of the previous mistakes were correct 6 weeks later. In this they surpassed the stronger group that worked with the teacher-produced transcript. Members of that group correctly produced only 47% of the previously noted and corrected items 6 weeks later.

Also, Mennim’s (2007) study of two Japanese students over 9 months found that their use of the word “garbage” became more accurate within that period. Notably, both students made long-term gains. Although I hesitate to draw conclusions for the longer-term effects of noticing from my results, I found that 95.5% (42 out of 44) of the items that were noticed and correctly produced in previous sessions were correct when they re-occurred in later sessions. This is a high rate, but because of the limited data available, and because it is difficult to know if these were corrections of systematic errors or mere slips that were performance-related (Lynch, 2007, p.316), this result should be cautiously interpreted.

**Research Question 3 and 4**

Research question 3 asks whether there are differences between learners who work alone and those who work collaboratively in terms of what they do with noticed items. As noted in Chapter 5, Individuals and Pairs behaved differently with respect to what they chose to notice, at what point in a session something was noticed, and how the item was treated after it was noticed. However, in the findings of this study in this area cannot be compared to the other transcription studies because none of them compared the noticing of learners working alone and those working in pairs. In this regard, this study is unique.
Research question 4 asks if the verbalization during noticing and its effects differ between learners working alone and those working collaboratively. Among the transcription studies, only Lynch (2001) and Mennim (2012) considered the collaboration between the learners, but neither study examined what the participants said as they collaborated and noticed items. As mentioned, Mennim (2012) focused on the strategies used in collaboration, but his study did not relate the strategies to the outcome of the noticing, unlike the present study.

Nonetheless, there is some basis for comparison between the present study and the ones reviewed with regard to collaboration. For example, participants in this study commented that one of the greatest frustrations working with a partner was when neither was sure of his or her hypothesis about an item under consideration, and lengthy discussions led nowhere. Mennim (2012), on the other hand, observed that his pair participants were more sanguine about impasses. He attributed this to their expectation that the teacher would be available to resolve the issue for them at a later time. The participants in this study did not receive teacher feedback as they proceeded through a Session.

As for the participants’ views on collaboration, Stillwell et al.’s (2010) study included information gathered from a survey administered to the participants. On that survey, the participants indicated that among the activities involved in the poster carousel task, of greatest value was self-transcription and self-correction, while other-transcription and correcting their partner were described as neutral or “not useful.” Conversely, the students thought the teacher’s feedback was “most useful,” which suggests that the students perceived peer input as less useful than direct teacher feedback. This result is contrary to what was found in the exit questionnaire for the present study. There the participants reported responding positively to collaborating with a peer to solve their language problems. In fact, the Pairs in the present study clearly indicated
that they recognized the value of pair work. However, this response was gathered at the end of partnerships that lasted over 6 Sessions. In working collaboratively, the Pairs often developed a shared sense of purpose. However, the partnerships depended on each individual’s willingness to collaborate. In the literature on collaborative dialogue (Swain and Watanabe 2013; Watanabe 2004; Storch 2002), researchers have found that some pairs collaborate better than others. Although the pooled knowledge of a pair is greater than each individual’s, the outcome is dependent on how well the individuals work together. As Lynch (2001) reported, one of his pairs was less successful at the task, possibly because of differences in gender and culture.

In the collaborative dialogue studies, there were aspects discussed that are consonant to this study. When participants verbalized their thoughts, it opened a window into how and what they were thinking. Talking was an integral part of the study. As noted in Swain (2006), however, not all talk has equal value. Verbalizing acted as a memory aid, but collaborative dialogue had greater value in focusing attention on particular difficulties. There was evidence that verbalizing drew attention to problems and heightened awareness compared to not talking at all. Items that were noticed but not verbalized were more often incorrect at the next attempt than items that were noticed and verbalized. This is similar to the results to some of the studies in verbalization and collaborative dialogue such as Qi and Lapkin (2001) and Tocalli-Beller and Swain (2005).

Differences were found between participants who worked alone and those who worked collaboratively. As mentioned in Chapter 6, Pairs evolved and sometimes developed a joint focus, or became increasingly focused on fluency and presentation strategies and skills. The Individuals’ behaviour remained relatively static for the duration of the study. The relationships built by Pairs were frequently mentioned in positive terms and described as having helped a
participant gain confidence, offer assurance, and give feedback as a partner in learning. This was also found in Shehadeh’s (2011) and Yarrow and Topping’s (2001) longitudinal studies.

In sum, while there are a number of results from this study that are similar to those of other studies, there are also differences. The differences in the results between the studies can be accounted for in terms of the design of the studies. Unlike most of the transcription studies, this study was not designed to fit within an existing curriculum or classroom. Consequently, in order to find out if students can mediate their own learning, no teacher input was provided, unlike the other transcription studies reviewed. In the contexts in which those studies were conducted, classrooms, it was only natural for the teachers in the studies to provide feedback to their learners.

Inspired by the transcription studies, the procedure of this study included verbalization at the transcript-noticing stage in an attempt to trace why something was noticed. As well, the design included opportunities to notice after every repeat of the performance. In order to trace development of language features and forms, the verbalization of the students as they collaborated to change and improve the presentation was examined. To determine if there were differences in verbalization between Individuals and Pairs, an Individual condition was included for comparison. In order to capture the changes among the participants over a period of time, the study was designed as a longitudinal study. Because of the differences between the transcription studies and the present study, the present study adds to our knowledge of noticing and differences in noticing between learners working alone and those working in pairs.
Section Three: Pedagogical Applications

**Transcription as a tool for self-assessment**

One pedagogical application is the use of transcription as a tool for self-assessment. Mennim (2012) noted that self-transcription is a pedagogically sound activity because it engages students. It also gives them a way to generate discussion and it encourages students to think about their own language use. He stated that the responsibility for L2 correction is not that of teachers alone, and L2 learners can take some responsibility for their own learning and use peers as a resource. This is aptly demonstrated in the findings of this study. Because transcription requires close listening, and produces a visible product, it is an excellent tool to highlight for learners the relationship between accurate grammar and pronunciation and conveying meaning. As well, it is a cognitively challenging task involving listening and writing, thus it has the potential to increase cognitive involvement with the product. The obvious drawback is that it is time-consuming. Also, some training is required to develop the ability to write exactly what was heard instead of what one thinks one heard or what is correct. Stillwell et al (2010) found that participants were fairly accurate in their transcription. This was also observed in this study. Participants became more accurate with practice over the Sessions.

**Audio and video self-recordings**

Repeatedly, I was struck by what my participants said during their exit interviews with regard to their participation in the study. The biggest surprise for many was hearing themselves speak English. It turned out that none of them had ever carefully listened to audio-recordings of themselves speaking the L2, and it was a discovery that jarred their perception of who they were and how they performed. A few commented that they disliked the quality of their voice during the Sessions: that it was too high-pitched or that they sounded silly speaking English. There were
participants who recounted how family members had reacted when they overheard their recordings while they did the transcription: from laughter to highly critical comments that were recanted to save face after a family member discovered that he or she was listening to the participant herself. For the participants, though, the initial reaction to listening to themselves was usually a mixture of surprise and dismay. It was a useful exercise for the students; they soon learned to use it as a diagnostic tool for understanding their own strengths and weaknesses.

When asked whether they would record themselves after the study as a way to monitor their performance and prepare for the public exam, two-thirds of the learners said they would.

This strongly suggests that students can benefit from hearing themselves speak English. Given the many devices that can make voice recordings today, from smart phones to laptops, which many students already use, students should be encouraged to record themselves and listen to the playback closely. It could illuminate the extent to which they are comprehensible to listeners. Voice recordings could also be used as models for students to use to practice articulation. Students could even shadow good models of English speakers. There is a wealth of Internet resources that could be used to this end. Thus, another pedagogical application is the use of audio self-recordings as a tool for teaching speaking in the classroom.

If the use of recording technology is embraced as a tool for the self-assessment of performance, a natural extension would be to video-record learners so that mannerisms and gestures while speaking English are captured in addition to spoken language. It is difficult to know how learners will react to seeing themselves speaking English on video. It is possible that some learners may not enjoy this. Lynch (2007) included viewing video playback of the participants’ performances, and some of the participants noted that it was “not useful” in the exit survey. But without more information, it is hard to know if whether it was not useful because it was distracting or because it discomforted the learner. With the proliferation of video technology
over the Internet for communication today, and the ease that students have using and consuming visual media, this is another pedagogical tool that could be used for examining performance for feedback and analysis.

Collaborative dialogue

A third application is the use of collaborative dialogue as a pedagogical tool in the language classroom. As mentioned, verbalizing, talking to one another serves many useful functions, such as rehearsal and memory aid, eliciting assistance, and offering suggestions. The opportunity to discuss in depth offers students the opportunity to engage and resolve problems for themselves. It gives them the chance to see the gaps in their knowledge. As noted in Chapter 2, a number of studies have found positive benefits for peer mediation, interaction, and co-constructed learning. As pairs work together, their collective histories and understanding become intertwined and their focus sometimes converges in particular areas that serve the purpose of the task. This collaboration often generates ideas or activates latent memories that help to produce a product or complete a task. In contrast, a student working alone does not have this resource.

At present, collaborative work is already often practiced in the form of group work in many classrooms in Hong Kong. However, a considered use of collaborative work that is extensive and sustained is not often seen in language classrooms. There may be obstacles, real or imagined, to adopting this kind of practice. Among the real ones are lack of time, large class sizes, the amount of noise generated, difficulty of supervising to ensure students are “on-task,” and fairness in assessment. The latter is because assessment is primarily based on individual work, and the idea of sharing marks from group work can be difficult to accept as fair for both teachers and students since competition is so keen. However, the chances of students engaging with a task and noticing the gaps in their knowledge, or seeing a bigger picture, are greater when
they work together than when they work alone within their own orbit and realm of experience. Of course, much depends on the individuals that make up the group, and the teacher’s comfort with noise and relinquishing control of the activity. But if the task is well-designed and challenging enough, the aims of the lesson could be better served by group work. In the long run, as teachers and students see the benefits accrued from mediated learning, the process could go more smoothly, as I have observed among the Pairs in this study.

**Strengthening and systematising the teaching of speaking**

This study led me to reflect on how students are taught to speak English in Hong Kong (Li, 2007; 2009). I believe there is room for English speaking pedagogy to be strengthened in the classroom. Chan and Li (2000) extensively documented the pronunciation problems of Hongkongers speaking English. Their paper is a good starting point for Hong Kong English language teachers seeking to understand the scope of their students’ difficulties. In a further series of papers, Chan (2006, 2007, 2009, 2010) examined the sources of these difficulties, many of which are due to first language interference. Chan and Li (2000) concluded with a suggestion that teachers prioritize the aspects of pronunciation to address depending on much they impede comprehensibility.

But even some common rules are worth teaching and reminding students to attend to. For instance, the pronunciation of words that end in “-ed” and “-es” depends on whether a consonant or vowel precedes them, and that the pronunciation of “the” depends on whether it precedes a word that begins with a consonant or a vowel sound are worth teaching. In general, students should also be made aware that speaking is different from writing. For example, hesitations and minor dysfluencies in spoken language are normal, and a pause is not a mistake. Taking time to
enunciate increases comprehensibility, while “sounding fluent” does not mean speaking at high speed.

Students also need to be aware that fluency and communicating the message are two different things. While both are at the heart of speaking well, an essential factor, which did not seem to be completely grasped by many of the participants, is that grammar is integral to meaning. Meanwhile, the greatest challenge for teachers remains how to create meaningful opportunities for students to use English to develop their confidence and fluency, given the limited use of English in everyday life in Hong Kong.

Section Three: Limitations and Directions for Future Research

There are several limitations to this study. As a small-scale study involving 24 students, the results cannot be generalized to a larger population, nor can the results of the study be taken to be predictive of learners of English in Hong Kong in the future. Although the study sheds light on what it is that learners notice of their speaking performance in this particular context, what learners notice in a different context cannot be determined from this study.

There are also limitations in terms of pedagogical applicability. Because of the iterative design of the study, it is not easy for teachers to implement these procedures in their classrooms, nor is it desirable. However, collaborative work has been shown in this study, and many others, (Storch, 2013) to be beneficial for student learning. Thus, adaptations of the procedure for collaborative tasks may be a productive route to take.

Another limitation is that the study design did not include opportunities to repeat performances using the exact same structures in order to find out if the gains made from noticing were durable.
To summarize, the study adds to our understanding of noticing speaking: what is noticed, how much is noticed, and what happens to that which has been noticed. Also, it adds to our understanding of how noticing alone is different from noticing in pairs. It illustrates the strengths of collaborative dialogue and peer mediation. Although the output between Individuals and Pairs was not dramatically different at PT 3, the composition process as well as the preparation for each attempt changed as a result of being in the Pair condition. This study also adds to our understanding of transcription as a tool for noticing speaking. It has proven to be an important tool for visualizing speech so that it can be inspected. The transcript becomes an artefact for the participants to reflect on and refine. The verbalization during the task of both the Individuals and the Pairs give us insight into what the participants were thinking and feeling as they approached their task.

Thus, as I conclude my study, I have no doubt that noticing is an important first step in learning. However, what follows is a complex set of choices that learners make in order to learn, and these choices are still not well understood. For teachers, the practical question remains how to teach something so that it is retained over the long term, so that students can produce grammatically correct utterances. There is reassurance from Mennim’s (2007) longer-term study, which concluded that long-term effects of noticing for the non-count noun “garbage” were positive, but that study involved only two students.

**Future Research**

One direction research could take is to examine frequency effects, task type as well as type of instruction, and whether noticing is self-generated or other-generated as it relates to speaking. Repetition, though it is thought to be the bane of classroom teaching, may have value for consolidating what has been noticed and learned. Since learning does not take place in a
vacuum, how learners engage must take into account their connections to their social, historical, and cultural setting. Thus, another direction for future research will lie in using a Vygotskian perspective to understand language development. The teaching method, the act of learning, who is learning and why, and by what means this is being pursued, all have a bearing on learner engagement. From the pursuit of studies from different but complementary perspectives, a fuller picture of how learning occurs could emerge to offer insights for researchers, teachers, and learners.

All things being equal, when a linguistic feature is presented explicitly or implicitly, why it is noticed at a particular time by someone, and then produced again within some range of accuracy remains somewhat of a mystery, no less because of the complex nature of the human mind, in each individual formed from a particular set of experiences. Nonetheless, I do not believe variations in individual conditions alone account for the differences, and I look forward to future research that discovers sound pedagogic practices that will foment noticing and bring about language acquisition.
REFERENCES


Hong Kong Examinations and Assessment Authority. (2002). *HKALE question papers: AS use of English*. Hong Kong, Hong Kong: Author.


Norwood, NJ: Ablex.


Language Teaching Research, 11(2), 143-161.


Appendix A

Sample Examination From the Use of English Paper D

Candidate A

Preparation Time (Total time: 10 minutes)

Part One – Individual Presentations (1½–2 minutes each; total time: 8 minutes)

Read the following passage. You may make notes to use as the basis for your presentation, but should make the presentation in your own words. If you wish, you can add your own opinions or ideas.

Net language

For years, researchers have been looking at the ways different people learn a second language and the findings are very interesting.

Joyce Chow, a Secondary Seven student, is a keen second language learner in Hong Kong who is constantly working hard to improve her English. She does not spend time on examination drills like many students do, nor does she attend tutorial classes. However, her English is well above average. Her secret to success is chatting on the Net.

Some people criticise Internet programmes like ICQ (I seek you) and Net Meeting because they waste time and encourage the use of sub-standard English. This ‘cyber-slang’ uses symbols and borrows grammar from other languages. For example ‘O3day so happy, ah?’ is acceptable use of English on the Net, although many people feel that this is ‘bad’ English. Joyce sees it differently – for her, the cyber world opens up many opportunities to make new friends that she can communicate with in English. But Joyce insists on the use of proper English with all those she meets, as a way to practise the language.

‘It’s an excellent way to use English and it’s convenient,’ says Joyce. ‘I can sit at home and learn by chatting to others. When I am chatting online, I need to respond quickly and I find that I can now think in English. It’s almost like speaking face-to-face.’ When Joyce comes across new words that she doesn’t understand, she stops for a while to look them up in the dictionary. She says she has picked up many new words that she can now use in her compositions at school. When she has time, Joyce reads through her old messages and reviews the words and phrases. ‘It trains my typing skills as well,’ adds Joyce. ‘Isn’t that much more interesting than exam practice?’

Part Two – Group Discussion (Total time: 2 minutes preparation, 10 minutes discussion)

Your school is organising a Language Week. As part of the activities, the English teachers at your school have asked your class to do a project entitled ‘Tips for Language Learners’.

You and your classmates have decided to produce a video for the project. You are meeting to discuss what information to include in the video, where and how to gather information and how the information could be presented in an interesting and attractive way.

You do not need to come to a final decision or to reach a conclusion in your discussion. You should try to discuss both the content and procedures involved in your discussion task. You cannot take notes or write during the discussion.

The examiners cannot answer any questions about the instructions or the task you have to discuss. If you have any questions, you may discuss them among yourselves and that will be an acceptable part of your group discussion, and will be assessed.
Candidate B

Preparation Time (Total time: 10 minutes)

Part One – Individual Presentations (1½ – 2 minutes each; total time: 8 minutes)

Read the following passage. You may make notes to use as the basis for your presentation, but should make the presentation in your own words. If you wish, you can add your own opinions or ideas.

Movie language

Like many Hong Kong students of his age, John Chan is a movie fan. He also speaks English fluently. The main reason for John’s success in English is the way he uses his hobby as a tool for language learning.

John loves watching movies of all kinds and he finds that he can learn a lot about vocabulary, expressions and accents from the characters in the movies. He claims that he never gets bored with watching films and can watch the same one up to ten times but still enjoy each and every part. ‘This helps me remember what I’ve learned in a better way,’ he explains.

When watching English movies, John always tries to listen to the actors as they speak and he only reads the Chinese subtitles when he has difficulty catching a word. He admits that he will happily watch a film over and over again, until he learns all the lines by heart. He then knows exactly what the actors are going to say in their next line so the next time he watches, he can say the words with them. Many teachers would say that memorisation is not an effective way of learning, but it works well for John.

Interestingly, John believes that watching Chinese movies is an equally good way of learning English and explains, ‘My mind automatically tries to translate what I hear from Chinese into English. And when I’m not too sure, I take a quick look at the English subtitles.’ This way John ensures he is always working on his English. He also looks at the subtitles when he wants to see how some common Chinese expressions are written in English. ‘Sometimes I like to guess how it will be translated in the subtitles, and then I check to see if I’m right,’ says John.

Part Two – Group Discussion (Total time: 2 minutes preparation, 10 minutes discussion)

Your school is organising a Language Week. As part of the activities, the English teachers at your school have asked your class to do a project entitled ‘Tips for Language Learners’.

You and your classmates have decided to produce a video for the project. You are meeting to discuss what information to include in the video, where and how to gather information and how the information could be presented in an interesting and attractive way.

You do not need to come to a final decision or to reach a conclusion to your discussion. You should try to discuss both the content and procedures involved in your discussion task. You cannot take notes or write during the discussion.

The examiners cannot answer any questions about the instructions or the task you have to discuss. If you have any questions, you may discuss them among yourselves and that will be an acceptable part of your group discussion, and will be assessed.

NOT TO BE TAKEN AWAY
Candidate C

Preparation Time  (Total time: 10 minutes)

Part One – Individual Presentations  (1½–2 minutes each; total time: 8 minutes)

Read the following passage. You may make notes to use as the basis for your presentation, but should make the presentation in your own words. If you wish, you can add your own opinions or ideas.

Local language

To most Hong Kong students, learning a second language and mastering it can be both difficult and boring. The vast majority of students will not feel secure until they have done lots of exam practice. They believe that by doing the most recent past papers over and over again, they will perform well in the public exams. But does this necessarily guarantee good results?

Victor Lee, a first-year University student, does not think so. His performance in A-level Use of English was excellent, but he did not bury himself in books and past papers before the examinations. He knows he is living in a bilingual city, so he prefers to ‘look out’ for English. This means he makes a particular effort to focus on the English used in signs, notices, advertisements and announcements. And he does it so often it has become a habit.

‘When I am shopping or simply walking along the street, I pay attention to information in English. This helps me to think in English,’ says Victor. ‘I know that many of my friends simply ignore things they see which are written in English like street names, brochures and posters, but I don’t. I enjoy learning new words this way.’ Victor usually takes along a notebook and jots down things he finds interesting. Then when he gets home he checks in his dictionary to see how the words are pronounced. Victor also loves eating out and reading the menu gives him the chance to discover names of dishes and drinks in English. ‘If I go to Western restaurants, for example in Stanley, I always order my food in English,’ he adds.

To Victor, learning English can be very interesting if you ‘look out’ for it. Hong Kong is a cosmopolitan city and there are many opportunities around for learning and practising English.

Part Two – Group Discussion  (Total time: 2 minutes preparation, 10 minutes discussion)

Your school is organizing a Language Week. As part of the activities, the English teachers at your school have asked your class to do a project entitled ‘Tips for Language Learners’.

You and your classmates have decided to produce a video for the project. You are meeting to discuss what information to include in the video, where and how to gather information and how the information could be presented in an interesting and attractive way.

You do not need to come to a final decision or to reach a conclusion to your discussion. You should try to discuss both the content and procedures involved in your discussion task. You cannot take notes or write during the discussion.

The examiners cannot answer any questions about the instructions or the task you have to discuss. If you have any questions, you may discuss them among yourselves and that will be an acceptable part of your group discussion, and will be assessed.

NOT TO BE TAKEN AWAY

31
Living language

Most English teachers believe that in order to master a language, a student must be given the opportunities to use it in real-life situations. Ingrid Leung, a seventh-former, is certainly a firm believer of this theory and she tries to bring English into all aspects of her life.

Ingrid is one of the most outstanding students in her school and she performs well in every subject. However, she gets the most satisfaction from being able to use and appreciate English. ‘I don’t see it as a foreign language – I see it as part of my life,’ Ingrid explains. At school, she will never miss a chance to talk to her teachers in English, especially the native-speaking English teacher. Sometimes, she will even initiate conversations with her schoolmates in English. She also likes to communicate with tourists she meets on the streets of Hong Kong, for example by giving them directions.

As a music lover, Ingrid enjoys listening to songs. In particular, she likes to hear English songs written by people of all nationalities, because she finds that writers from other parts of the world use English differently. She often reads the lyrics and admires the sounds of the language as she sings along. She also tries to see any stage productions of English plays and musicals that are shown in Hong Kong. This summer she is taking part in her school’s English drama production and she will play a leading role.

‘My ambition is to study at university and then to become an English teacher,’ says Ingrid. ‘English is important because it is a world language, and I want to encourage my students to learn it in a natural and enjoyable way.’ In Ingrid’s opinion, learning English is much easier and much more fun if you make it a part of your life.
Appendix B
Pilot Study

Inspired by the literature, I decided to investigate the connection between transcribing as a route to noticing and speaking. What did learners notice of their spoken English? What would they change if they could say something again? Would talking things through, alone or with a friend, be helpful in deciding how to say it again? After developing a procedure for data collection, I conducted a pilot study. I recruited 6 Hong Kong ethnic Chinese students who were on exchange for a term at major Canadian universities. They were in their second or third year of study, and their majors were in business and law, government and law, commerce and linguistics. I recruited them because they approximated the age and experience of senior high school English learners. In addition, all of them were familiar with the Use of English Paper D (Oral) examination prompts that I was using to collect spoken data, since they had done this as part of their university entrance examinations. I met each of these students once. Two students met me individually, while the remaining 4 met me as 2 pairs. At each meeting, we conducted only Part A (individual presentation) of the examination, which was recorded. The participants transcribed this afterwards immediately on their own or together as a pair. They were told to look over the transcript they produced, and if necessary, make changes to improve it alone, or with the help of the peer partner. Next, they were asked to do the presentation again, which was again recorded. This second recording was played back immediately, and they were asked to listen carefully to it, stopping whenever they wanted to comment on any changes they had made, or failed to make, and anything else that came to mind.

I analysed the following: the student’s transcription of their first presentation and compared this with my transcription of both their presentations; the comments they made after each attempt at the presentation; and I noted the changes they made between recording 1 and 2. I
found that they were all able to render a generally accurate transcription of their own recording, although there were a few differences between my version and theirs. One of the participants noted that she found herself writing what ‘should be’ rather than what she actually heard. There were also instances where the participants had written down the past-tense verbs that ended with ‘–ed,’ or the plural ‘-es’ form for a noun, where I did not hear the /dɪd/ or /ɪz/ suffix at the end of the words. Overall though, they were able to transcribe their own presentations with little difficulty, and fairly quickly.

Overall, the pilot study showed that the participants noticed a range of things in their spoken presentation, which included grammar errors, pronunciation difficulties, sentence structure and organization of the presentation, and the dysfluencies in speaking. While they were able to make target-like self-initiated corrections, they were not able to identify every mistake they had made. Most of the time, what they had discussed and talked over while they prepared the transcription were acted upon and taken up in the second presentation. It was not clear whether pair or individual work led to more noticing. For pairs, it seemed to depend on the dynamics and relationship between the two. While one pair worked together in a productive fashion, the other pair was reticent with each other and hardly gave each other any helpful feedback. Meanwhile, participants who worked alone were focused on what they had self-identified and carried out the changes in the second presentation.

All the participants also reported a heightened awareness of the way they spoke. For the second presentation, they all claimed they spoke more slowly. One participant said that he was thinking carefully about each word he was uttering to ensure the grammar was correct. Another said she had less ‘flow’ as a result. Two participants thought that both presentations were about the same, while one was disappointed that she did not produce a perfect presentation the second time around, though she admitted that perhaps her expectations were too high. One participant
became aware of the speech fillers he used, and consciously avoided saying them in the second presentation, dramatically reducing these by 85%. In all, the participants claimed that this awareness would make them pay more attention when they spoke English in the future. Though they reported that they were aware of the mistakes they made even as they spoke, getting their message across was more important than producing a grammatically accurate delivery.

With the pilot completed, I reformulated the research questions and refined the procedure for data collection. Noticing was operationalized in three ways: transcribing their own two-minute oral presentations; comparing transcripts of the same performance done by themselves with classmates’ and discussing these differences, either alone or with a peer; and by listening to their own oral presentations immediately after their attempt to improve the performance.
Appendix C
Questionnaire for Participants at the Start of the Study

Choose a nickname that you wish to be identified as: ....................................................................................

Age: ........................................ At what age did you start learning English?.............

Which language do you speak most frequently
  i) at home? ......................................................................................................................................................
  ii) in the classroom?........................................................................................................................................
  iii) during group projects with classmates? ....................................................................................................
  iv) when socialising with friends?...................................................................................................................

Which language(s) do you use most when you are:
  i) thinking through a problem / working on school work ................................................................................
  ii) thinking about your friends (relationships etc) ...........................................................................................
  iii) making important decisions .......................................................................................................................
  iv) thinking about things in general ................................................................................................................
  v) writing e-mails ............................................................................................................................................

Put a tick in the column that best represents your opinion on the following statements:

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 I enjoy learning English in school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 My spoken English is better than my written English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 I like speaking English in school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 I enjoy preparing for the HKAL English Paper D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 I don’t like English, but I have to learn it for the sake of my future</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 I am satisfied with my level of written English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 I have positive experiences using English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 I try to find opportunities outside school to use English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Given a choice, I always prefer using Chinese over English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 If there is a choice, I always prefer watching a movie in Chinese</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Other people can understand me when I speak English</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 I am confident in using English for conducting business (non-social conversations) over the phone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please feel free to write any other comments you have on your language use in the space on the next page.
Thank you!

Thank you!
Appendix D
Participant Questionnaire at the Completion of Study

(The questionnaire was spaced over 3 pages to provide adequate space for participants to write their comments.)

Your nickname: .......................................................

1. How did you find the additional oral practice sessions? Did you find them useful or enjoyable?

2. Please comment on the transcribing process.
   a) What did you learn from transcribing yourself?
   b) What did you learn from transcribing your partner or other people in the study?

3a) Do you think your English speaking skills have improved over these weeks? Why or why not?

3b) If you think you have improved, can you name three things specifically about your speaking that you have changed?

4. Has participating in this study affected your level of confidence in presenting in English? If yes, how have you been affected?

5. Besides speaking, has participating in this study had an impact on other aspects of your English skills (e.g. reading, writing, listening)?

6a) If you worked with a partner: (if not go to 6b),
   i) What was the best thing about working with your partner?
   ii) What was the worst thing about working with your partner?
   iii) Do you think you would have come to the same conclusions / arrived at the same solutions if you had worked alone?
   iv) How did you feel about thinking aloud with a partner?

6b) If you worked alone:
   i) Do you think you would have benefitted from working with a friend?
   ii) How did you feel about talking aloud as you worked through the transcript comparison?

7. After participating in this study, are there areas that you have targeted specifically for yourself to work on? Please specify.

8. Do you think you will record and transcribe yourself as you prepare for the HKAL exams or other oral presentations? Why or why not?

9. Do you think you will “talk aloud” when you are explaining things to yourself or working on challenging questions? Why or why not?
   a) If yes, what language will you be using as you talk aloud?

Please feel free to write any other comments you have on your language use in the space below.

Thank you very much for taking part in this study!
Your participation is very much appreciated ☺
Appendix E

Summary of Data for Questionnaire Administered at the Beginning of the Study

Average age of participants when they first started learning English: 3.79 years
Appendix F
Summary of Data for Questionnaire Administered at the Beginning of the Study

Responses for Participant Questionnaire 1

<table>
<thead>
<tr>
<th>Question</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) I enjoy learning English in school</td>
<td></td>
</tr>
<tr>
<td>Disagree, 4</td>
<td>Agree, 19</td>
</tr>
<tr>
<td>Strongly agree, 1</td>
<td></td>
</tr>
<tr>
<td>(2) My spoken English is better than my written English</td>
<td></td>
</tr>
<tr>
<td>Disagree, 6</td>
<td>Agree, 16</td>
</tr>
<tr>
<td>Strongly agree, 1</td>
<td></td>
</tr>
<tr>
<td>(3) I like speaking English in school</td>
<td></td>
</tr>
<tr>
<td>Disagree, 9</td>
<td>Agree, 15</td>
</tr>
<tr>
<td>Strongly agree, 0</td>
<td></td>
</tr>
<tr>
<td>(4) I enjoy preparing for the HKAL English Paper D</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 2</td>
<td></td>
</tr>
<tr>
<td>Disagree, 6</td>
<td>Agree, 19</td>
</tr>
<tr>
<td>Strongly agree, 3</td>
<td></td>
</tr>
<tr>
<td>(5) I don't like English, but I have to learn it for the sake of my future</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 5</td>
<td></td>
</tr>
<tr>
<td>Disagree, 7</td>
<td>Agree, 9</td>
</tr>
<tr>
<td>Strongly agree, 3</td>
<td></td>
</tr>
<tr>
<td>(6) I am satisfied with my level of written English</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 4</td>
<td></td>
</tr>
<tr>
<td>Disagree, 16</td>
<td></td>
</tr>
<tr>
<td>Strongly agree, 4</td>
<td></td>
</tr>
<tr>
<td>(7) I have positive experiences using English</td>
<td></td>
</tr>
<tr>
<td>Strongly agree, 2</td>
<td></td>
</tr>
<tr>
<td>(8) I try to find opportunities outside school to use English</td>
<td></td>
</tr>
<tr>
<td>Strongly agree, 1</td>
<td></td>
</tr>
<tr>
<td>Disagree, 12</td>
<td>Agree, 11</td>
</tr>
<tr>
<td>(9) Given a choice, I always prefer using Chinese over English</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 1</td>
<td></td>
</tr>
<tr>
<td>Disagree, 3</td>
<td>Agree, 12</td>
</tr>
<tr>
<td>Strongly agree, 8</td>
<td></td>
</tr>
<tr>
<td>(10) If there is a choice, I always prefer watching a movie in Chinese</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 3</td>
<td></td>
</tr>
<tr>
<td>Disagree, 15</td>
<td></td>
</tr>
<tr>
<td>Strongly agree, 6</td>
<td></td>
</tr>
<tr>
<td>(11) Other people can understand me when I speak English</td>
<td></td>
</tr>
<tr>
<td>Disagree, 1</td>
<td>Agree, 22</td>
</tr>
<tr>
<td>Strongly agree, 1</td>
<td></td>
</tr>
<tr>
<td>(12) I am confident in using English for conducting business (non-social conversations) over the phone</td>
<td></td>
</tr>
<tr>
<td>Strongly disagree, 1</td>
<td></td>
</tr>
<tr>
<td>Disagree, 9</td>
<td>Agree, 12</td>
</tr>
<tr>
<td>Strongly agree, 2</td>
<td></td>
</tr>
</tbody>
</table>
**Appendix G**

Summary of Data for Questionnaire Administered at the End of the Study

<table>
<thead>
<tr>
<th>Question</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) How did you find the additional oral practice sessions?</td>
<td>24</td>
<td>0</td>
</tr>
<tr>
<td>(2) What did you learn from transcribing yourself?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Became more aware</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Pronunciation</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Dysfluency</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Rate of speech</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>(2b) What did you learn from transcribing others?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Style</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Common mistakes</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Pronunciation</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Rate of speech</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Dysfluency</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>(3a) Do you think your speaking skills have improved?</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>(3b) Can you name three specific things that have changed?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Rate of speech</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Pronunciation</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Tenses / grammar</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>(4) Has this study affected your level of confidence?</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>(5) Has this study had an impact on other aspects of your English skills?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6a) Best thing about working with a partner</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Share / learn from one another</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>(6b) Worst thing about working with a partner</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unable to resolve a problem</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>(6c) Would you have come to the same conclusions?</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>(6d) How did you feel talking with your partner?</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>(6bi) (For Individuals) Would you have liked to work with a friend?</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>(6bii) (For Individuals) How did you feel talking out loud?</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>(7) Are there specific areas that you have targeted to work on after the study?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Pronunciation</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>(8) Will you record yourself to practice for the exam?</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Transcribe</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>(9) Will you talk aloud to explain things to yourself? Which language?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depends on the subject</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Cantonese</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>
Appendix H

Code Book

Column 2: C

There are 2 numbers coded for each item in the C column. The number on the left refers to the type of item: 1 represents Grammar, 2 - Pronunciation, 3 - Content, 4 - General comment, 5 - Difference in transcription. The second number identifies the session, either 1, 2 or 3.

Examples: 11 – Grammar item, first session
32 – Pronunciation item, third session
41 – General comment, second session
51 – Difference in transcription, first session

Column 3: Code

There are two sets of information here. The first identifies the one who mentions the item first. There are 3 possibilities – Self (S), Other (O), or Both (B). For B, it usually applies to the student transcripts, where both the participant under focus and the other transcriber identified the same item for attention.

The second set of information is related to the category of the item.

Examples:
B-tense - Both identified, wrong tense
S-agree – Self identified, subject-verb-agreement
O-verb - Other identified, verb form
O-accept - Other identified, mispronunciation
S-discourse - Self identified, arrangement of information

A detailed description for each category is given below.

1 Grammar

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>adverb</td>
<td>adverbs, adjectives, -ly, adverbials</td>
</tr>
<tr>
<td>agree</td>
<td>subject-verb agreement, verb+s; has/have; is/are</td>
</tr>
<tr>
<td>article</td>
<td>Determiners – a, an, the; this, that, these, those; some, any</td>
</tr>
<tr>
<td>gender</td>
<td>using the correct gender</td>
</tr>
<tr>
<td>modal</td>
<td>will/would; shall/should; can/could</td>
</tr>
<tr>
<td></td>
<td>e.g. she cans borrow books</td>
</tr>
<tr>
<td>noun</td>
<td>±-s or ±-es</td>
</tr>
<tr>
<td>phrasal verb</td>
<td>e.g. keen on vs. keen in</td>
</tr>
<tr>
<td>pronoun</td>
<td>using the correct form: he, him, she, her, this, these, that, those</td>
</tr>
<tr>
<td>sentence</td>
<td>changing the way a sentence is phrased (adding or deleting words),</td>
</tr>
<tr>
<td>structure</td>
<td>retaining same meaning</td>
</tr>
<tr>
<td>tense</td>
<td>±d, ±ed; irregular verb forms</td>
</tr>
</tbody>
</table>
| compa | comparatives and superlatives  
|       | e.g. more happier, more easier |
| verb  | verb form: ± ing;  
|       | e.g. was sink vs. was sunk |

### 2 Pronunciation

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>accept</td>
<td>accepts that the pronunciation is inaccurate</td>
</tr>
<tr>
<td>deny</td>
<td>denies mispronouncing word; attributes difference in transcription to other student mishearing</td>
</tr>
</tbody>
</table>

### 3 Content

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
</table>
| info    | ± information; incorrect / inaccurate information.  
|         | e.g. this is not needed, or I added a point  
|         | e.g. it should be EU, not UN                                                                 |
| meaning / lexis | changing the meaning; using a more appropriate word to express intended meaning; or not know for certain what a word should be  
|         | e.g. She can take money and save money. Ya, it is better to use, withdraw and deposit. |
| discourse | order of information presented; introduction; conclusion; cohesion: repetitive linking devices e.g. using and to begin most sentences. |

### 4 General comment

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>improved</td>
<td>improvement observed</td>
</tr>
<tr>
<td>fluent</td>
<td>fluency or lack of it observed</td>
</tr>
<tr>
<td>no change</td>
<td>No change in the quality of the performance</td>
</tr>
<tr>
<td>opinion</td>
<td>presence or absence of an opinion; the need to include and opinion</td>
</tr>
<tr>
<td>filler</td>
<td>observing the presence of fillers, hesitations, pauses</td>
</tr>
<tr>
<td>worse</td>
<td>the performance is worse than the previous one</td>
</tr>
</tbody>
</table>

Other comments about nervousness, forgetting, lacking confidence, speech speed and volume are also coded under general comment

### 5 Difference in transcription

A difference in the transcription is noted either by the focus participant or the partner in the session. The differences are usually minor, relating to transcription differences in placement of pauses, different ways of representing dysfluencies and typographical differences.
Appendix I
Coding Rules

There are 11 columns in the grid used for coding. Column 1 is for numbering, while Column 2 codes for category of item noticed, and the session in which it occurs. The third column identifies the person who noticed the item first, and gives more detailed information on the type of item. What is first identified is what is coded as, even if subsequently, it becomes a different type of item (e.g. first noticed as an error in the tense, but subsequently used as part of a modal auxiliary.)

The remaining columns contain excerpts from the student transcripts and the transcript of the session to track what happens to the item that is noticed. An item is coded as being noticed at 4 points in the process – either as the student does the transcription of the exam practice, during the transcript comparison stage, during the stimulated recalls of the second attempt or third attempt (Columns 4, 5, 8 and 11).

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>C</td>
<td>Code</td>
<td>Transcript</td>
<td>Noticing transcript</td>
<td>Production Task 2</td>
<td>Correct?</td>
<td>SR1</td>
<td>Production Task 3</td>
<td>Correct?</td>
<td>SR2</td>
</tr>
</tbody>
</table>

To follow the item through the process (Columns 4 to 11), the excerpted transcripts are further coded. The coding rules are listed below:

First code: this applies to Columns 4 (Transcript) and 5 (Noticing Transcript).

a. If a change was suggested in the student transcript (Column 4), but nothing was said during the transcript comparison session, then Column 4 is coded. Depending on the suggestion, Column 4 is coded either:
   | C = correct, or
   | I = incorrect.

b. If something was mentioned about the item during the transcript comparison session, then Column 5 is coded:
   - C – if they arrived at a correct solution
   - I – if they arrived at an incorrect solution
   - Q – if they note the problem but have no answer
   - NS – if they suggested a tentative answer for which neither are sure
   - N – if they simply stated that there was a problem, but no further discussion took place.

c. If the item was not used in the exam practice and so, did not appear in the student transcript, nor was mentioned in the comparison of the transcripts, then Column 5 is coded as –.

Second code: this applies to Columns 6 (Production Task 2) and 7 (Correct?).

a. The columns are coded:
   - C – if the change made was correct
ii) **ANCI** – accept, no change, incorrect: if the student had previously considered a change necessary, but made none and the result was incorrect

iii) **ANCC** - accept, no change, correct: if the student had previously considered a change necessary, made it, and the result was correct

iv) **DNCI** – disagree, no change, incorrect: if the student had disagreed with the suggested change, made no change, and the result was incorrect

v) **DNCC** – disagree, no change, correct: if the student had disagreed with the suggested change, made no change, but the result was correct

vi) **CSC** – changed structure, correct: used a different structure or form that was not discussed, and the result was correct

vii) **CSI** – changed structure, incorrect: used a different structure or form that was not discussed, but the result is incorrect

viii) **CMTL** – correct, more target-like: made a change, but the result was not really correct, only more target-like

ix) **CLTL** – correct, less target-like: made a change, but the result was even less target-like than the original, but not completely incorrect

b. If the structure is not used at all in PT 2, it will be coded **D** – disappeared, in SR1.

**Third code: this applies to Column 8 (Stimulated Recall 1).**

a. If nothing was said about PT 2, then Column 8 is coded **-**.

b. If the student noticed something about PT 2 during SR1, Column 8 is coded:
   i) **C** – if they arrived at a correct solution, or noted that the item was correct in PT 2
   ii) **I** – if they arrived at an incorrect solution
   iii) **Q** – if they note the problem but have no answer
   iv) **NS** – if they suggested a tentative answer for which neither are sure
   v) **N** – if they simply stated that there was a problem, but no further discussion took place.

**Fourth code: this applies to Columns 9 (Production Task 3) and 10 (Correct?).**

The codes used are the same as those in the Second Code (Columns 6 and 7)

**Fifth code: this applies to Columns 11 (Stimulated Recall 2).**

If the student noticed something about PT 3 during SR2, Column 11 is coded:

i) **C** – if they noted that the item was correct in PT 3
   ii) **I** – if they arrived at an incorrect solution
   iii) **Q** – if they note the problem but have no answer
   iv) **NS** – if they suggested a tentative answer for which neither are sure
   v) **N** – if they simply stated that there was a problem, but no further discussion took place.

b. In addition, if the student noticed something about PT 3 in SR 2, and provides a correct solution, it will be coded **D3** – Disappeared, 3<sup>rd</sup> attempt, because we do not know if it will be correctly produced in future. This code is only used in this column.
General comments
These are not coded, but categorised and counted as things the students mention.

Difference in transcription
Likewise, these are also not coded, because they do not recur. The excerpts from the transcripts show that the students noticed:

i) differences in orthographical representation of dysfluencies
ii) typographical errors
iii) punctuation differences
iv) pauses in different places

These are also not deeply processed. If they deny or debate them, then they are categorized under Grammar or Pronunciation. Nonetheless, these are counted because they give an indication of what is being noticed, and reveals something about the participant in terms of attention to detail (albeit 'spot-the-difference' noticing), or lack of attention to detail. Missing the opportunity to see the difference as a possibility of their mispronunciations could be a product of over-confidence, and could indicate shallow processing.
Appendix J
Coding for Engagement When Verbalizing

1 Cursory
(Tracy 73, T3, S-verb)

Transcription
In my opinion, I think that if we cannot afford those clothes, just trying (try) to wear it

Transcript Noticing
J: trying?
T: what did I?
J: here, just try, and trying

2 Verbalized intended change

(Sandra Poon 4, T1, O-article)

Transcription
I have just read a article

Transcript Noticing
First I see a change here that he mm hai she heard that I said ah article rather than she correct me is ‘an’ article … I read I really write ‘a’ article. … is … I need to use ‘an’

3 Explanation / Meta-language

(Phoebe Leung 6, T1, B-preposition)

Transcription
when she was um in (on) the MTR,

Transcript Noticing
P: When she was on the MTR, she was in the MTR. I think she is inside the MTR show I will I will use the word ‘in’ the preposition ‘in’ the MTR

... P: And, ya, in the MTR vs ON the MTR, I think is IN R*: in is more correct
P: ya. Because inside the MTR, in the MTR.

4 Unsure but makes a choice (∓ consult passage)

(Karina 48, T2, O-info)

Transcription
G: Recently, during her in her exhibition of wildlife of photos, she has collected 20 thousands of signatures, those people are against the government who er actually who sell the land to some developers that the land was actually for (?) the country park.

Transcript Noticing
G: .... Did you mention, why, did she, collect, the signatures?
K: No. Uhuh, can I take a look at that? (the original passage)
G: Ae, ok ah, I think but tell specifically what, what the signature is about. [the girls try
to figure out what the ‘country park’ point is]
K: (reads softly from original) to sell land which had previously been set aside for
country parks

5 Only at SR2

(Purina 46, T2, O-modal)
Transcription

and she finds out (found) that the websites have can remember your preference of
choosing web pages,

Only verbalized at SR2
C: can remembers
P: (laughs)
R*: So what should it be?
C & P: can remember!

R* is the researcher. All other letters represent the participants.
Appendix K
Difficulties Encountered During Data Collection

There were a few difficulties encountered during data collection, one of which was scheduling. There were constant changes to the school schedule at the school, so the participants and I were frequently re-negotiating meeting times. For instance, I might find on arrival at the school that due to preparations for a major school event, the school administration had declared a shortened time-table for the day, or had given senior students time off regular lessons to help with preparations, or were using Monday’s schedule on a Thursday to make up for lost Mondays. Sometimes, the students had class tests scheduled during lunch-time. There were also occasions when lab experiments that were begun the period before lunch had to be completed during lunch time. These resulted in frequent juggling of the data collection schedule. With a lot of accommodation from the participants, I was able to complete data collection over the course of the school term.

Another difficulty was finding suitable rooms to collect data in. Although the school welcomed me, they were short on space and many of the places that I was allocated were not ideal for recording. For instance, the room we frequently used was by the main entrance of the school and the sound of buses roaring by occasionally drowned out the student, as did other students traipsing noisily along during lunch, or bouncing basketballs across the hall. On a few occasions, the participants and I were displaced when the room was needed for emergency meetings with parents or was turned into a temporary storage room. We improvised by doing recordings in a science lab, a room next to where the choir was practicing, or open spaces in the school on occasion. Nonetheless, the participants were highly motivated, accommodating and flexible. Some of them even had to lunch while they worked, because that was the only time they could meet. In spite of some of these difficulties, most of the data are clearly audible and usable.
Appendix L

An Example of an Extended Discussion

Froggie and Dull

R: so it’s just one part that is the problem? (3.0) Any other points?
D: the conclusion part
R: conclusion part. What’s wrong with the conclusion?
D: what’s wrong? Um very wrong uh (laughs)
R: why is it very wrong?
D: uh
F: cause she’d, focusing, on, learning English instead of read and write
D: but, what I would like to say is, um, to learn English well, we have to, read more and write more, that, this a way to suggest, how, apart from what, affect carries case, we can also read and write more in, when learning English
F: but if I just, read the last sentence, I, ae don’t know what, you’re talking about
D: uh ya, yes
R: mm-hm
D: I don’t really well I’m talking- what I would like to, uh the audience won’t know, what I’d like- what message I would like to convey in the last part
R: mm-hm. So how would you change this?
D: um, actually I want to say, ae, ae to learn good English is not, something that can be, done in just a very short period but we have to, um, we have to, the whole process has to be, has to be, carried out, in continue- ae continuously.
R: mm-hm.
D: ya. I wou- I want to say this but then (laughs) I don’t know why I end up, ended up, like this (taps the page)
19:38
D: aiya. What is my opinion ah?
F: your opinion?
D: hai-ya
F: we can learn English everywhere. Wei we have to practice, more.
D: ae
F: you ask me, your opinion
D: uh I, what I said just now can I, can I, put it in a conclusion part?
F: you said
D: I said, I said ae, mm, what did I say just now? Ae, what did I say just now? Uh, oohhh, said (laughs out loud)

F: we can (laughs)

D: what did I say just now ah?

F: you said we, can learn, mhai we learn English

D: uh uh English, can’t be improved within a short period,

F: of time

D: of time but instead, we have to, it is a continuous process. We have to keep on, keep on

F: improving

D: keep on practicing, uh, hai mai leh?

F: keep on learning new words and keep on, el-ae,

D: keep on learning, ae,

D: Ae, In my opinion, I think (laughs), I think ae, work, English, can’t be improved, within a short period of time. It’s, uh is uh, the whole process, will have to last, the whole process has to be carried out, carried out continuously. Uh-hohh.

F: what process?

D: mm, has to--

F: you repeat your conclusion again

D: I think that, English, can’t be improved within short period, short period of time. We have to, read and write continuously.

F: (laughs) God! Everyday, every second we are, we have to

D: If we want to make a improve re- we must

F: read more and write more

D: and write more--. Uh. Full-stop. Mm. Ok.

28:35

D: um, the conclusion part is still, still, can be improved I think

R: did you get to say what you wanted to say?

D: no

R: no. what did you want to say?

D: I have told, ah ah ah, what I want to say ah? We have to keep on,

F: learning

D: |

R: | (garbled) doing with your hands?

F: uh we-

D: (laughs)
F: ae we have to, ae, ae you mean we j- you don’t just learn English in lessons but also in our daily lives like things around us
D: uh-huh
F: we learn from, things around us. Ae, ya is that what you?
D: uh maybe maybe um, be-cusing, uh English, isn’t something that can be, improved within a short period of time so that, we have to,
F: we have to--
D: blah blah blah what she said
F: we have to, grab our chances
D: (laughs)
F: to learn English, whenever we can
D: necessary
F: xxx, and throughout our lives, our life (laughs)
D: aiya
R: is that what you’re trying to say?
30:03
F: uh, not 100%
D: let me think
F: mm hai-lo. Think, what do you think ah?
D: drink more.
30:29
D: what do you think? What can I add, in ah? How to improve xxx
F: ae, English can’t be improved within within a short period of time
D: so that, if you want improvement,
F: If we, hai-lor
D: we must, we must,
F: we must
D: we must
F: if we want to have good English,
D: uh-huh
F: not only improve. We want to have good English. We must, ae--, learn, English
D: learn mmhai we must write and, think ae mhai wo, learn and
F: we must, sp-
D: speak English in our day to day life. Aiya. So stupid, are to piece ah. AhhhH!!
F: we must ae, we must try to learn, English in, hh-hh, at any time, anywhere.
D: we must
R: forgot, to set 4 minutes
Four minutes begin at 31:52
D: we must learn and and we must learn
F: we must learn, aiya, so stupid gah!
D: me too ahh!
Both girls laugh (they are energized, and work at a quicker pace)
D: if we want, if we want our English to be improved, we have to, we have to
F: learn, we have-
D: we have to take the initiative to learn and write more, so that, ae, instead of, instead of receiving ae, aiya meh- passively ah, aiyah!!
F: (laughs)
D: instead of acting passively we must take the initiative to learn and write
F: initiative to learn
D: more often so that, as a result, our English can be improved.
F: so?
D: so?
32:39
D: in my opinion, I think that English can’t be improved within a short period of time. Instead of, instead of acting passively, we must take the initiative to learn and write,
F: like like what Victor Lee did (laughs)
Both girls laugh.
D: what Victor Lee does ah, hmai-ah?
F: ae hai-uh
D: uh. Ae we must read, we must take the initiative to read and write more often, like Victor Lee does. As we result, our English, our English--
F: ae Victor Lee does everyday.(laughs)
D: does everyday ok ok.

PT2
In my opinion, I think, English can’t be improved, within, short, period of time. If we want to make a proo- improvement, we have to, read- more- and- write- more.

PT3
In my opinion, I think that English can’t be improved, within a short period of time. Instead of acting pacsif- passively, we should, take the initiative, to-- read- more- and- write- more like, Victor, Victor Lee does, everyday.
# Appendix M

## Glossary

Abbreviations and terms used in the thesis:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIT</td>
<td>Difference in transcription</td>
</tr>
<tr>
<td>SR 1</td>
<td>Stimulated Recall 1 (includes transcribing at home)</td>
</tr>
<tr>
<td>SR 2</td>
<td>Stimulated Recall 2</td>
</tr>
<tr>
<td>SR 3</td>
<td>Stimulated Recall 3</td>
</tr>
<tr>
<td>PT 1</td>
<td>Production Task 1</td>
</tr>
<tr>
<td>PT 2</td>
<td>Production Task 2</td>
</tr>
<tr>
<td>JP</td>
<td>Judgement Possible</td>
</tr>
<tr>
<td>NJP</td>
<td>No judgement possible</td>
</tr>
<tr>
<td>Trackable</td>
<td>A unit of analysis containing a point of First Noticing and a Product(s)</td>
</tr>
<tr>
<td>Untrackable</td>
<td>Items that do not contain a point of First Noticing and a Product</td>
</tr>
<tr>
<td>FN</td>
<td>First Noticing</td>
</tr>
</tbody>
</table>