EXPLORING THE POTENTIAL OF REHEARSAL VIA AUTOMATIZED STRUCTURED
TASKS VERSUS FACE-TO-FACE PAIR WORK TO FACILITATE PRAGMATIC AND
ORAL DEVELOPMENT

By

Tetyana V. Sydorenko

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Second Language Studies

2011

ABSTRACT

EXPLORING THE POTENTIAL OF REHEARSAL VIA AUTOMATIZED STRUCTURED TASKS VERSUS FACE-TO-FACE PAIR WORK TO FACILITATE PRAGMATIC AND ORAL DEVELOPMENT

By

Tetyana V. Sydorenko

Past research has uncovered ways to improve communicative competence, including task-based learner-learner interaction (e.g., R. Ellis, 2003) and task planning (e.g., Mochizuki & Ortega, 2008). Teacher-guided planning particularly increases the benefits of learner-learner interaction (Foster & Skehan, 1999). One component of communicative competence – pragmatics – requires special attention as it may be one of the most difficult aspects of a second language (L2) (Blum-Kulka & Sheffer, 1993).

Judd (1999) had proposed that guided/structured practice of pragmatics should take place before open-ended production. To investigate this claim, this classroom-based, semester-long quasi-experimental study examines whether planning via computer-delivered automatized structured tasks (CASTs) can help develop communicative competence, including pragmatics. The research questions are: (1) What is the effect of task planning (via CASTs, in pairs with classmates, or no planning) on pragmatic appropriateness and fluency, accuracy, and complexity of oral responses? (2) How does this effect transfer to new tasks? (3) What are learners' attitudes towards CASTs and pair-work activities?

Three groups of ESL learners participated in the study. The *structured task* group (*N*=17) engaged in planning via CASTs. Learners in the *pair-work* group (*N*=16) planned the tasks via role-plays in dyads. The control group (*N*=17) did not engage in planning. The tasks were role-plays focusing on pragmatic appropriateness. The specific type of planning that learners engaged

in was rehearsal. The immediate effect of rehearsal was measured via gains on rehearsed tasks, while the effect of rehearsal on language development in general was measured via gains on new tasks at the end of the study.

The results of mixed-design ANOVAs indicate that all groups improved on both rehearsed and new tasks only in terms of vocabulary diversity. All groups also improved on pragmatics on new but not on rehearsed tasks. First, these findings suggest that when performing speech acts, learners focus on improving informational content rather than fluency, accuracy, and syntactic complexity of oral responses. Second, pragmatic development may not be apparent immediately.

In this study, the experimental groups did not differ on gains for any of the measures. If there are differences between rehearsal via CASTs versus rehearsal via learner-learner dyads, the use of a more interactive computer program that allows for the provision of feedback after rehearsal might be necessary to uncover them. While the experimental groups had the same learning outcomes, their discourse patterns during task rehearsal differed.

The results do not provide conclusive evidence that task rehearsal is beneficial because the control group did not differ from the experimental groups on gains. However, the relative shortness of pragmatic instruction, lack of immediate feedback, access to pragmatics in the second language environment, and low proficiency level of the learners may have lessened the effect of task rehearsal on pragmatic development.

The questionnaire data indicate that while 57% of respondents enjoyed both CASTs and pair work, most students favored one type of activity or the other partly due to their personal preferences. The findings are discussed in terms of implications for using CASTs in face-to-face, hybrid, or online courses.

DEDICATION

To my family and friends for their unconditional love and support.

ACKNOWLEDGEMENTS

The faculty, staff, and colleagues in the SLS program and the English Language Center also deserve to be acknowledged for being supportive every step of the way. While all of them have left their mark, my special thanks goes to Dr. Debra Hardison for spending hours on writing letters of recommendation. I also greatly appreciate countless suggestions from my colleagues Maren Schierloh, Ching-Ni Hsieh, Amy Thompson-Öztan, Angelika Kraemer, Shaofeng Li, Luke Plonsky, and Chien-Hsiung (Scott) Chiu throughout the dissertation process. I am equally grateful to Joan Reid, Joy Campbell, Alicia Rice, Carol Wilson-Duffy, Pat Walters, and Dr. Dan Reed, who have made the administrative aspects of this process seamless.

The development of authentic materials constituted a significant part of this study. For generously donating their time (and acting skills) to help create authentic materials that the students in this study enjoyed immensely, and in some cases for rating the data, I extend my sincere gratitude to my friends and colleagues: Ben White, Luke Plonsky, Sara Hillman, Jennifer Behney, Mark Shea, Kari Richards, Ann Letson, Becky Swab, Peter Hoffman, Sally Behrenwald, Hyesun Lee, Mariah Shafer, Rebecca Floyd, Chad Eakin, Mary Frances Cappiello, Marie Hollenbeck, Sheila Wheeler, Allyssa Chamberlain, Carson Maynard, Gad Lim, Aaron Ohlrogge, Sarah Goodwin, Caitlin Gdowski, Avis Jackson, and Micah Cousins; as well as to Alyse Bedell, Katie Murphy, Amanda Jacobs, and Nathan Bynum (undergraduate students at MSU). I would also like to thank Mica Fanning and Dave Ragan for taking interest in these materials and adopting them in their courses as part of the study. Finally, I greatly appreciate the tremendous contribution of ESL students at MSU, who agreed to participate in the study and to offer their feedback on the novel program they were asked to use.

Last but not least, I would like to thank my beloved family and friends for preparing me for this path, believing in me and, in some cases, for inquiring regularly and tirelessly how many

pages have been written ☺ Mom, dad, Irisha, Dean, mom Kim, dad Patrick, Geraldine, Dominic, and all my dear friends - your moral support has been priceless! And special admiration goes to my husband Adrian for never-ending patience and endurance.

With all my heart – Thank You!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Purpose of the Study

This study addresses the central question of how to develop second language learners'

pragmatic competence and, consequently, their communicative competence. Specifically, the

aim of this study is to investigate whether performance of communicative pragmatically-focused

tasks can be effectively planned via computer-delivered automatized structured tasks (CASTs).

Research suggests that task planning improves subsequent performance on a task (e.g., Crookes

1989; Foster & Skehan, 1996; Guará-Tavares, 2008; Mehnert, 1998; Ortega, 1999; Sangarun,

2005; Skehan & Foster, 1997) and that task planning is most efficient when guided by a teacher,

as opposed to individual or group planning (Foster & Skehan, 1999). This may be due to the fact

that language modeled by a teacher is more correct and richer than what the learners can offer

each other, especially at lower proficiency levels. In a similar vein, Judd (1999) had proposed

that in pragmatics instruction, learners should engage in guided/structured practice before they

perform open-ended tasks. However, while instructors can guide a class as a whole, guiding each

student individually is difficult, to say the least, in large classes. Individualized instruction is

often advantageous because it allows learners to work at the pace comfortable for them. Such

individualized guidance may be possible through CASTs. In these tasks, designed specifically

for the present study, learners are presented with a sequence of video-recorded prompts related to

a pragmatic situation, and learners are instructed to respond to these prompts in a manner that is

pragmatically appropriate in a given situation. The prompts can be video-recorded by native

speakers (NSs), in which case learners can receive the benefits of both NS input and guided

planning. Learners can rehearse their performance by completing these structured tasks several

times. Because CASTs can be completed outside of class, they may be beneficial for preparing learners for language production in class.

Since learners will be practicing speaking as they perform pragmatically-focused tasks in oral modality, gains in pragmatic appropriateness, as well as fluency, accuracy, and complexity of oral responses will be examined. That is, both pragmatic and oral development will be evaluated.

1.2 Rationale for the Study

One of the major goals of foreign or second language instruction is to raise language learners' communicative competence (e.g., Savignon, 1997, 2001). A variety of communicative competence models have been developed (e.g., Bachman, 1990; Bachman & Palmer, 2010; Canale, 1983; Canale & Swain, 1980; Celce-Murcia, 2008; Celce-Murcia, Dörnyei, & Thurell, 1995; Martínez-Flor, Usó-Juan, & Alcón Soler, 2006). While these models differ in how communicative competence is divided into components, all of them include the knowledge and ability to use the sound system, vocabulary, and linguistic structures of a language appropriately and effectively, as well as communicate in pragmatically-appropriate ways. Pragmatics involves the knowledge of the proper linguistic forms (pragmalinguistics), as well as an understanding of when to use them (sociopragmatics) based on such factors as power, social distance, and degree of imposition (Kasper & Roever, 2005). Cultural knowledge is also a part of communicative competence (Celce-Murcia) because culture determines which behaviors are pragmatically appropriate and which ones are not.

In the field of second language acquisition (SLA), continuous progress is being made towards methods of improving learners' communicative competence; however, pragmatics is still in need of special attention as it may be one of the most difficult aspects of a second

language (L2) (Blum-Kulka & Sheffer, 1993). Research has shown that obstacles to advanced proficiency include misunderstanding culturally appropriate ways of expressing gratitude, refusing an invitation, apologizing, complimenting, and dealing linguistically with appropriate forms of regret, demands, and indirectness (Biesenbach-Lucas, 2007; Cohen & Shively, 2007; Kondo, 2008; Lantolf & Thorne, 2006; Rose & Ng, 2001; Sykes, 2008). Sociopragmatics is generally more difficult to acquire than pragmalinguistics (Roever, 2009). While exposure to the target language in a second language setting contributes to the development of sociopragmatic competence more than learning in a foreign language environment (e.g., Bardovi-Harlig & Dörnyei, 1998; Cook, 2001; Kondo, 1997; Matsumura, 2001; Niezgoda & Roever, 2001; Rose, 2000; Shardakova, 2005), it is still not sufficient (Bardovi-Harlig & Hartford, 1993; Wilkinson, 2002; Winke & Teng, 2010). Research has shown that although learners obtain part of their pragmatic knowledge "for free" through positive transfer and universal principles of politeness and communication (Kasper & Rose, 2001, p. 6), various aspects of pragmatics need to be taught explicitly (Bardovi-Harlig, 2001; Jeon & Kaya, 2006; Kasper & Rose, 2001; Rose, 2005) in order to raise learners' awareness of certain strategies and forms that are different from those in their first language and to speed up their pragmatic development. For example, language learners need to be taught how to structure discourse in a target language (Holmes, 2003; Young, 1994).

Research shows that pragmatics is amenable to explicit instruction (see Roever, 2009, for a review). Unfortunately, pragmatic instruction is often not incorporated into the syllabus due to the lack of class time (Kasper & Schmidt, 1996) or little support in course design (Bardovi-Harlig, 2001). Yu (2008) found that in communicatively based classes, EFL teachers in Taiwan did not provide instruction on pragmatics although it would have been appropriate to do so when learners displayed gaps in sociopragmatic competence. Even when pragmatics is taught in class,

textbook materials are often not authentic and do not provide appropriate input (Bardovi-Harlig, 2001; LoCastro, 2003; Usó-Juan, 2008a, 2008b; Vellenga, 2004) (but see Feak, Reinhart, & Rohlck, 2009, for materials designed by conducting conversation analysis of natural data). Thus, Martínez-Flor and Usó-Juan (2010) call for the development of authentic materials for teaching pragmatics, including a variety of multimedia.

To raise language learners' communicative competence, including pragmatics, researchers and teachers advocate the use of task-based instruction (e.g., R. Ellis, 2003; Skehan, 2003), which involves asking students to perform meaningful tasks using the target language. A large part of research in the area of task-based instruction is on task planning because "Practically it can help inform the methodology of task-based teaching, where one of the options available for implementing tasks concerns whether or not to allow students time to plan and, if so, what kind of planning and for what length of time" (R. Ellis, 2009, p. 474). Task planning can be conducted in a form of pre-task strategic planning (i.e., planning for but not actually performing a task), rehearsal (performance of a task several times), or online planning (i.e., thinking about the task as one performs it). Studies examining the effect of different kinds of planning on speech production have found that, depending on the type of planning, fluency, accuracy, complexity, or a combination of these aspects improve (see R. Ellis, 2009, for a comprehensive summary). It has been also found that strategic planning is most beneficial when guided by an instructor, followed by learners' individual planning, and then followed by planning in groups (Foster & Skehan, 1999). However, task-planning research needs to be diversified. For example, only one task-planning study has examined the development of pragmatic competence (Taguchi, 2007). The effects of guided planning, such as guidance provided by an instructor, versus unguided planning have been also understudied (Foster, 1996;

4

Foster & Skehan, 1999; Mochizuki & Ortega, 2008; Skehan & Foster, 2005). Additionally, the majority of pre-task planning studies have been on strategic planning rather than rehearsal. To fill several major gaps in second language research described above (i.e., the need to focus on pragmatic instruction and the lack of studies on task planning involving pragmatics, guided planning, and rehearsal), the main goal of the present study is to compare the effect of task rehearsal on pragmatic and oral development in two conditions: rehearsal via CASTs and rehearsal in learner-learner dyads.

1.3 Overview

In chapter 2, I will review the relevant literature on the development of pragmatics and speaking skills. This will include existing models of teaching pragmatics and speaking and the theoretical bases for them. I will then propose ways in which pragmatics and speaking instruction can be enhanced by CASTs. The methodology of task design, data collection, and data analysis are explained in chapter 3. The results are provided in chapter 4 and discussed in detail in chapter 5. Both quantitative analyses of learning gains and qualitative analyses of discourse patterns and learner attitudes were conducted. This dissertation concludes with chapter 6 on implications for the teaching of pragmatics and speaking via CASTs and limitations and future directions of such research. In this study, the investigation of CASTs took place in face-to-face English as a second language courses. However, implications for online and hybrid/blended language courses are also discussed.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the relevant literature on the development of pragmatics and speaking skills. The existing models of teaching pragmatics and speaking are outlined, and the theoretical basis is provided for each of them. I will then propose ways in which pragmatics and speaking instruction can be enhanced by CASTs. The chapter concludes with the research questions for the current study.

2.1 Pragmatics and Speaking Instruction

The consensus among researchers is that pragmatics should be taught explicitly (Bardovi-Harlig, 2001; Jeon & Kaya, 2006; Kasper & Rose, 2001, 2002; Rose, 2005). The reason for this view is that even after lengthy exposure to an L2 in study-abroad or target language contexts, learners do not fully acquire pragmatic forms and context-appropriate meanings (Bardovi-Harlig & Hartford, 1993; Wilkinson, 2002; Winke & Teng, 2010). According to Kasper and Rose (2001), while learners obtain part of their pragmatic knowledge through positive transfer and universal principles of politeness and communication, many aspects of pragmatics need to be incorporated into instructed language settings in order to raise learners' awareness of certain strategies and forms that are different from those in their first language. This follows Schmidt's (1993) contention that relevant pragmatics features might go unnoticed by learners even with prolonged exposure to the language and that attention to "linguistic forms, functional meanings, and the relevant contextual features" is necessary for pragmatic learning to occur (p. 35). Schmidt (2001) postulated the *noticing hypothesis*, the strong version of which states that "while there is subliminal perception, there is no subliminal learning" (p. 26). This claim is based on numerous studies which found that previously well-learned but unattended information is registered in the memory for a brief period of time, but no learning occurs without focal attention

or awareness. Schmidt (2001) argues that many features of L2 input are non-salient, and thus "intentionally focused attention may be a practical (though not theoretical) necessity for successful language learning" (p. 23). Explicit instruction is thus necessary to point out such non-salient features to the learners (N. Ellis, 2011). While grammatical features, especially those that carry little meaning, are often considered to be the least salient (see Schmidt, 2001), pragmatics also has non-salient aspects (e.g., Kasper & Rose, 2002; Schmidt, 1993).

Another reason why pragmatics should be taught explicitly is that learners generally do not receive feedback on the pragmatic appropriateness of their statements from natural interactions in the L2 (S. Gass, personal communication, July 19, 2011). For example, expert users and even learners of a language may draw speakers' attention to their incorrect use of vocabulary and grammar, but when learners make pragmatically inappropriate statements, they are generally assumed to be rude rather than incorrect. This happens because pragmatics is closely tied to social norms, while other aspects of the language rarely have that connection. Although learners' pragmatically inappropriate statements may have consequences, for example, they may not be invited to a party by the speaker they were inadvertently rude to, learners may not know what exactly led to these consequences. However, during pragmatics instruction in a language class, learners can be explicitly told which statements are inappropriate in specific situations and why.

In research on pragmatics instruction, speech acts have been studied the most (Bardovi-Harlig, 2002; Kasper, 2006; Kasper & Rose, 1999; Roever, 2011) and will be the focus of this study. Speech acts were originally proposed by Austin (1962) and Searle (1969), and they can be characterized as communicative functions (e.g., Cohen, 1996; Sykes, 2008), such as apologizing, requesting, or complimenting. Speech acts can be realized in different ways, depending on the

context. For example, depending on the severity of the situation, appropriate apologies could range from "Sorry" to "I'm so sorry! Is there anything I can do?" Following Blum-Kulka, House, and Kasper (1989), speech act realizations will be referred to as strategies in this study, although they have been also referred to as semantic formulas (e.g., Bardovi-Harlig, 2001). Finally, several strategies are often needed to achieve one communicative function, in which case a speech act sequence would be performed (Blum-Kulka, 1997).

Several models of speech act instruction have been proposed. Olshtain and Cohen's (1991) five steps of speech act instruction are: (1) diagnostic assessment, (2) model dialogue, (3) evaluation of the situation, (4) role-play activities, and (5) feedback and discussion. Diagnostic assessment can be conducted in a number of ways, but perhaps the fastest way to get a broad picture of learners' knowledge of speech acts in an L2 is by using multiple-choice questions. Learners are provided with situations and dialogs, and they are then asked to choose a pragmatically appropriate continuation of each dialog from a list of available options. In Olshtain and Cohen's model, the second step is the provision of a model dialog, the purpose of which is to expose learners to authentic input as it is vital for pragmatic development (Bardovi-Harlig, 2001). Model dialogs can be shown as videos of native speakers performing targeted speech acts, and these dialogs should be as naturalistic as possible. However, students should not only be exposed to input, but they should also evaluate the situations in model dialogs, which could be done as both pre-listening and post-listening activities. This goes along with the idea stated above that many pragmatic features are not salient, and thus learners' attention should be explicitly drawn to input. Via this process, learners can begin to develop their sociopragmatic competence. Next, students should engage in the communicative practice of speech acts via role-play activities, which promotes automaticity, fluency, and draws learners' attention to forms (see section 2.2 for

functions of output). The final stage is feedback and student-teacher discussion, during which learners can compare speech act behavior in the target language and in their L1 and become aware of the similarities and differences in pragmatics between the two languages.

Judd's (1999) model is similar to the one by Olshtain and Cohen (1991), but it does not include diagnostic assessment and it is less specific. Following Judd, the teaching of speech acts can be broadly conceived as the following sequence: (1) cognitive awareness, (2) receptive skill development, (3) and productive use. Learners' awareness can be raised by both explicit instruction and student discovery via questionnaires, interviews, and observations. As part of explicit instruction, students should be provided with detailed information on "participants, their status, the situations, and the speech events that are occurring, because merely presenting linguistic formulas without such background information can lead to overgeneralizations on how speech acts function in real-life situations" (p. 155). Judd also suggests that "learners should act as amateur ethnographers and collect their own examples of speech acts" because recordings of role-plays and self-reports of native speakers may not be entirely accurate and naturalistic (p. 156). However, this approach also has its own weaknesses as data collection by learners may be time consuming, learners of low proficiency level may not recognize speech acts in the target language environment, and learners may not have access to native speakers from a variety of backgrounds. The next phase proposed by Judd is receptive skills development. That is, learners rely on the information they obtained from the cognitive awareness stage to recognize a variety of speech acts, their organization, linguistic forms associated with particular speech acts, and sociopragmatic factors that affect the choice of pragmatic strategies. During this stage, a large amount of input should be provided so that students can contrast different forms and meanings. Finally, in the third stage learners can engage in productive use, first through structured teacher-

guided tasks, such as cloze, and then through open production tasks, such as role-plays and debates. The main point Judd is making about the third stage is that learners should first engage in structured practice so that their production is focused and their attention is drawn to specific forms, while open practice should be meaningful, communicative, and naturalistic.

Félix-Brasdefer's (2006) model echoes those of Judd (1999) and Olshtain and Cohen (1991) and includes the following: (1) identifying speech acts in audio or video input and raising cross-cultural awareness; (2) engaging learners in conversation analysis to draw their attention to input; and (3) providing opportunities for communicative practice (e.g., role-plays between students) and feedback. Félix-Brasdefer argues that engaging learners in conversation analysis is necessary in that students need to recognize and learn how to produce openings and closings, how to organize their discourse, and when to take turns.

In the three models of speech act instruction reviewed above, role-play is a commonly used task for practice (e.g., House, 1996; Koike & Pearson, 2005; Kondo, 2008; Liddicoat & Crozet, 2001; Sykes, 2005, 2008; Tateyama, Kasper, Mui, Tay, & Thananart, 1997; Trosborg, 1995) because, as Willis (2004) states, role-playing a real-life situation comes close to an authentic task in a classroom environment. Judd (1999), however, proposes that controlled / structured practice of pragmatics should take place before open-ended production of such tasks as role-plays. Similarly, Martínez-Flor and Usó-Juan (2006) suggest the following activity for a controlled / structured practice of pragmatics: while students watch a pragmatically-focused video, an instructor pauses it periodically and asks the students to guess what will be said next in the video.

Cohen's (2005) approach of using strategies in learning speech acts complements the three above-mentioned models. These strategies focus on (a) what learners can do to gather

information on various speech acts, (b) how speech acts can be practiced, and (c) how learners can pre-plan and monitor their production of speech acts. For example, Cohen suggests that learners should first identify which speech acts to learn based on their frequency of occurrence and their value in communication. Then learners should gather information on how speech acts are performed in a particular speech community via such means as observations, interviews, and even by asking native speakers to model performance in particular situations. Next, learners should perform a cross-cultural analysis of the sociopragmatic norms, pragmatic strategies, and linguistic forms used in a particular speech act. Once learners are ready for practice, they should seek a variety of opportunities to try out their production of speech acts in both imaginary and real situations with fellow learners and native speakers. Additionally, learners should ask for feedback from native speakers. When producing speech acts, learners need to use metacognitive strategies to decide how much pre-planning of the speech act they need to do beforehand, how they will be monitoring their production, and how they will evaluate their performance. By raising learners' awareness of various strategies that can be used independently from classroom instruction, Cohen's approach encourages autonomous learning of pragmatics.

Regarding speaking, task-based instruction is a frequently used and studied approach, and it can be used to improve communicative competence (Bygate, Skehan, & Swain, 2001; R. Ellis, 2003; Skehan, 2003; Wesche & Skehan, 2002). Task-based instruction can involve three stages: pre-task, during task, and post-task (R. Ellis, 2006). The following definition of a *task* will be used in this study: "an activity, susceptible to brief or extended pedagogic intervention, which requires learners to use language, with emphasis on meaning, to attain an objective" (Bygate et al., 2001, p. 11). The aforementioned approaches for teaching pragmatics have similarities with task-based instruction as they involve (1) pre-task stage for providing input and raising

11

awareness, (2) during task stage, usually a role-play, and (3) post-task stage for providing

feedback. That is, pragmatic instruction and speaking instruction can be easily integrated.

2.2 Learner Interaction

Task-based instruction is informed by the interactionist approach (R. Ellis, 2003; Skehan,

2003), which finds interaction to be paramount for language learning. As Gass and Mackey note

(2006, p. 176), "The interaction approach attempts to account for learning through the learners'

exposure to language, production of language, and feedback on that production." In this

approach, input, interaction, and output all play an important role in language learning. The

significance of input was proposed by Krashen (1982, 1985). However, evidence gathered from

French immersion programs in Canada indicated that abundant and comprehensible input is not

sufficient for language acquisition. As a reaction to this finding, Swain (1985, 1995, 2005)

postulated the *Output Hypothesis*, which states that language production (via both speaking and

writing) and modification of that production are also necessary processes in language acquisition.

At the very least, output promotes automaticity (Gass & Mackey, 2006; Segalowitz, 2003) and

fluency (Swain, 2005) through practice. Swain (1995) outlined three other functions of output.

The noticing/triggering function is that learners may notice what they do not know how to

convey as they are attempting to produce the L2. In other words, through the process of

production, learners may consciously recognize some of their linguistic problems. As postulated

in the *noticing hypothesis* mentioned above, noticing is a prerequisite for acquiring the target

forms (Schmidt, 1990, 1993, 1994, 1995, 2001). That is, while it is important that learners

process language for meaning, second languages cannot be fully acquired if attention to form

does not take place. The second function of the output is that it may be a "trial run" for learners

as they are testing their hypothesis of how a particular form may work in the L2 (Swain, 2005, p.

476). Finally, the metalinguistic (reflective) function of the output is that second language learning is also mediated by learners' reflection on their production or the production of others. This reflection can be verbalized. Thus, it was found that when working on a language task together, learners can produce language-related episodes (LREs), defined by Swain and Lapkin (1995) as any part of a dialogue in which students talk about the language they are producing. Learners' collaborative metalinguistic discussion and subsequent modification of output have been shown to enhance language development. According to the interaction approach, interaction is a third necessary component in language acquisition. In the process of interaction, learners are not only exposed to input (i.e., positive evidence) and practice the L2 through production, but they can also obtain negative evidence via feedback in a form of negotiation of meaning or overt corrections. This allows learners to notice gaps in their L2 knowledge and modify their linguistic system in an L2 (Gass, 1997; Long, 1996). That is, learners may notice even more gaps through interaction and feedback rather than through production alone.[1]

While many researchers and practitioners believe that interaction between learners is important, some have expressed concerns that learner-learner interaction might not be quite as effective because learners can incorporate each other's errors. However, Bruton and Samuda (1980), Gass and Varonis (1994), Machado (2000), and Ohta (2001) found that such cases are very infrequent. Thus, Ohta suggested that benefits of learners working together far outweigh the dangers. A similar concern is that sometimes students' misunderstandings of each other go unnoticed (Kobayashi, 2003). It is important for learners to notice and point out communication breakdowns because learners need to identify gaps in their knowledge in order to improve. As noted by Kim and McDonough (2008), communication problems are more likely to not be

---

[1] While interaction is also referred to in the Output Hypothesis, it is not central to this hypothesis.

resolved in pairs with low-proficiency learners. Proficiency has been found to play a role in interaction, with higher-proficiency learners producing more LREs than low-proficiency learners (Leeser, 2004; Williams, 1999). Watanabe and Swain (2007) found that it is an interaction between learners' proficiency and group dynamics that affects language development: in collaborative groups, both lower and higher ability learners can gain benefits from interaction.

Given that the development of language abilities requires communication in the L2, willingness to communicate is a part of the process (Dörnyei, 2003). A variety of factors can affect learners' attitudes towards interaction with peers in a foreign language. De Saint Léger and Storch (2009) found that learners' willingness to communicate increased as they became more confident in their L2 proficiency. Learners of similar proficiency may feel more comfortable interacting with each other, while those whose proficiency is lower than that of their interlocutors often feel insecure in group work (S.-J. Kang, 2005; Kowal & Swain, 1997). Additionally, working with learners of the same L1 makes students less eager to communicate (S.-J. Kang). This is understandable because such communication feels unnatural, and in many language classrooms with homogeneous student populations in terms of L1 this is a valid concern. S.-J. Kang and Cao and Philp (2006) also found that smaller rather than larger groups, interlocutor familiarity, and topic familiarity made participants feel more secure about interaction. In summary, while students can help each other learn, various factors can reduce the effectiveness of their interaction.

2.3 Task Planning

Preparation for an activity is one way to improve the effectiveness of learner interaction. Pre-task planning as a form of preparation improves subsequent performance on a task (e.g., Crookes 1989; Foster & Skehan, 1996; Guará-Tavares, 2008; Mehnert, 1998; Ortega, 1999;

14

Sangarun, 2005; Skehan & Foster, 1997) because attentional burden is reduced (R. Ellis, 2005).

Foster & Skehan (1999) also argue that "providing learners with greater planning opportunities should have a beneficial effect on the course of language development, since planned second language discourse should push learners to extend what they are capable of saying" (p. 218).

Task planning has been categorized as pre-task planning (i.e., preparation for an activity) and online or during-task planning (i.e., thinking about a task as one performs it) (R. Ellis, 2005). Pre-task planning is further divided into strategic planning (planning but not performing a task) and rehearsal (performing the same task several times). Different types of planning improve different aspects of oral performance, and planning studies have generally examined accuracy, complexity, and fluency. For strategic planning, there seems to be a more consistent positive effect on fluency than on accuracy or complexity. Specifically, most learners focus on fluency, and then some choose to also focus on accuracy and others on complexity (R. Ellis, 2009). Following Skehan's (1996) proposal, it is likely that during strategic planning, learners have to choose which aspects they will focus on due to their limited processing capacity. Learners' proficiency level may also be a significant factor (R. Ellis, 2009). More controlled studies are needed before conclusions are drawn about the effects of strategic planning. For online planning, that is, when learners are given as much time as they need to perform a task, it was found that complexity and accuracy increased, but fluency did not (Yuan & Ellis, 2003).

Due to a small number of studies on task rehearsal and their varying designs, it is not entirely clear what aspects of oral performance task rehearsal has an effect on. In Gass, Mackey, Fernandez, & Alvarez-Torres (1999), participants showed improvement in overall proficiency, selected morphosyntax, and lexical sophistication. In Bygate (2001), there was an increase in fluency and syntactic complexity, but not accuracy. Other studies have examined the effect of

task rehearsal on the information structure of the message. Bygate and Samuda (2005) found that a subsequent performance of the same task increased the complexity of learners' discourse, as shown by the amount of framing (i.e., any language additional to the narrative content of a story, such as characters' feelings or previews to what is about to happen). Németh and Kormos (2001) investigated the effect of task rehearsal on argumentation. They concluded that "in the repeated version of the task, familiarity with the task structure helped learners pay more attention to the informational content of their message, which was reflected in the higher number of supportive moves they produced" (p. 213). In a case study of two learners, Lynch and Maclean (2000) found that when learners were not told what to focus on, rehearsal improved different aspects of performance. In their study, learners participated in a poster carousel task, in which they modified their poster presentations based on the feedback from classmates who asked questions about the posters. The learner with low proficiency improved on accuracy, while the learner with high proficiency improved mainly on the clarity of the message, specifically on the explanation of a complex concept. That is, the nature of the task and the specific gaps in learners' knowledge may have an effect on the kinds of changes that will occur after task rehearsal.

Another important finding is that rehearsal improves subsequent performance on the same task (Bygate, 2001; Bygate & Samuda, 2005; Gass et al., 1999; Lynch & Maclean, 2000, 2001; Németh & Kormos, 2001), but not necessarily on a new task of the same format (Bygate, 2001; Gass et al., 1999). While there is no clear evidence that task rehearsal facilitates second language acquisition, the effects of rehearsal used over a prolonged period of time on new tasks have not been investigated. Bygate (2001) suggests that intensive use of rehearsal might contribute to language development.

While there are numerous studies on task planning, few planning studies have been conducted on pragmatics instruction, even though Cohen (1996, 2008) has already stressed the need of planning specifically for pragmatics, both strategic planning and rehearsal. As a pre-task strategic planning activity, several studies have examined the effect of consciousness-raising on pragmatic development (House, 1996; Rose & Ng, 2001; Takahashi, 2001; Takimoto, 2009; Tateyama et al., 1997). Taguchi (2007) examined the effect of length of learners' strategic planning time before performing a role-play on their pragmatic and oral development. However, no studies have examined rehearsal as a type of planning for a pragmatic task. Since in pragmatics instruction, role-play is a commonly used task, learners could rehearse a role-play several times as a way of preparing for the same role-play with a different learner. In fact, this is often done in language classes, although the outcomes of such an instructional technique have not been reported in empirical research. Some teachers may feel that rehearsal of the same task is boring for the students, but Bygate and Samuda (2005) state that if learners are encouraged to work differently on the same material, they will benefit from such rehearsal and be interested in it. That is, during the first enactment of the task, learners can create a rough draft of their performance, while the second time they can elaborate on it.

While task performance generally improves after planning, participants involved in the planning process can make a difference in the outcome. Foster and Skehan (1999) found that strategic planning offered most benefits for accuracy, fluency, and complexity when done with the teacher, followed by individual planning; group planning was the least effective. They suggest that controlled (or guided) planning is the key because it is standardized and prepared by the teacher, thus having a greater degree of organization. Mochizuki and Ortega (2008) and Skehan and Foster (2005) found that guided planners as opposed to unguided planners focused

17

less on fluency and more on accuracy, complexity, or both during task performance. Guided planning seems to be especially beneficial for cognitively demanding tasks (Foster, 1996).

2.4 Computer-Delivered Automatized Structured Tasks (CASTs)

Given that task planning has been found to improve learners' subsequent performance, some have proposed that learners should prepare for a task before class so that class time can be used more efficiently (Zyzik & Polio, 2008). This proposal may be relevant especially for pragmatics because there is generally lack of class time to be devoted to this aspect of a language (Kasper & Schmidt, 1996). Task planning is more effective when learners are guided through the process by a teacher (Foster & Skehan, 1999). However, since a teacher would not be present during task planning outside of class, it would be advantageous if learners are guided by other means. In this study, computer-delivered automatized structured tasks (CASTs) are used to investigate whether learners' task planning has an effect on language learning and, therefore, whether task planning could effectively take place out of class. While, to my knowledge, task planning via CASTs has not been investigated before, Judd (1999) has suggested that in pragmatic instruction controlled / structured practice, such as practice via a cloze task, should precede open production of role-plays.

One kind of a cloze task is a discourse completion task (DCT). A DCT is a dialog on a social situation in which a response of another person has been deleted, and students have to fill the blank according to how they think they would respond in a real-life situation (e.g., Blum-Kulka et al., 1989). The situation and the context are provided. DCTs could be written or oral, and they can have one or multiple blanks. When there are multiple blanks, there are also rejoinders: utterances from an imaginary interlocutor that follow the blanks. In this study,

18

CASTs are like multi-turn oral DCTs, except learners cannot preview the rejoinders beforehand. That is, learners only hear the next rejoinder after they respond.

DCTs have been used extensively for assessing pragmatic knowledge (see Roever, 2008, for a review); however, their potential for pragmatic instruction has not been investigated. DCTs could be conceptualized as structured tasks because learners need to provide responses to rejoinders in a particular order. These tasks may be useful in several ways. Via structured tasks, learners could observe the discourse organization of pragmatic routines. NS input could be also provided through such structured tasks. When performing such tasks several times, learners may notice, and perhaps even begin using, new words and phrases from the input. NS input is necessary for learning all aspects of an L2 (e.g., N. Ellis, 2009; Gass, 1997; Gass & Mackey, 2006; Krashen, 1982, 1985), but learners especially lack authentic input on pragmatics (Bardovi-Harlig, 2001) because there is a paucity of authentic materials for teaching pragmatics (Bardovi-Harlig, 2001; LoCastro, 2003; Martínez-Flor & Usó-Juan, 2010; Usó-Juan, 2008a, 2008b; Vellenga, 2004) and because exposure to a variety of speakers and situations is necessary especially for this aspect of a language where there is vast individual variability (Sykes, 2008). Although learners can receive authentic input via other means, such as by watching videos of NS interactions, learners might be more engaged with the input and notice more linguistic forms when they also have to respond to the rejoinders (or, in other words, prompts) in CASTs.

While ideally learners would be exposed to authentic input by talking directly to native speakers, learners at lower proficiency levels may be apprehensive of such interactions. When investigating anxiety and speaking, Woodrow (2006) found that learners were most anxious when talking to native speakers. Some of the reasons mentioned by learners were that they would not be understood by native speakers or that native speakers would not find it interesting to talk

to them. However, Dewaele, Petrides and Furnham (2008) found that a higher level of self-perceived proficiency in a language is associated with lower communicative anxiety. Thus it is possible that after completing CASTs, learners may be more prepared and therefore feel less anxious when talking to each other in class and to native speakers in the target language environment. These tasks may also provide a low-risk environment for self-paced practice, which is consistently mentioned as one of the benefits of computer-assisted instruction (Abras & Sunshine, 2008; Chan & Kim, 2004; Garrett, 1991; Levy, 2009; Sykes, 2008). For example, lower proficiency learners may need more time to plan a task than higher proficiency learners.

Another benefit of using CASTs is that learners would be talking less to interlocutors of the same L1, which is unavoidable in foreign language classes and even, at times, in second language classes. When interlocutors are of the same L1, they can switch to L1 rather than use the L2 all the time.

Given that this study is exploratory, it will not be investigated whether CASTs are beneficial because they provide opportunities for structured rehearsal, authentic input, or both. However, at a minimum, we can say that authentic input provided through CASTs should be beneficial because learners generally lack access to it and because it has been clearly attested in prior research that authentic input is crucial for language learning. Future research will be needed to disentangle the variables of structured rehearsal and authentic input.

Following various models of pragmatics instruction and task-based instruction, consciousness-raising and explicit instruction should precede production, and feedback should be given after production. In this study, CASTs are conceived of as one kind of activity that learners can do as a form of task planning, specifically task rehearsal. That is, the computer program used in this study to deliver CASTs currently does not provide explicit instruction to the learners and

it does not have the capability to provide feedback on learners' performance. Instead, consciousness-raising activities can be led by an instructor to provide initial explicit pragmatics instruction before learners engage in rehearsal. After self-paced rehearsal via CASTs, learners can first engage in rehearsal with each other in class and then receive feedback from an instructor in class.

2.6 Research Questions

In this study, the effect of task rehearsal in two different ways is examined. The tasks are pragmatic speech acts. The three conditions are (1) rehearsing given situations individually via CASTs, (2) rehearsing given situations through role-plays in learner-learner pairs, or (3) no opportunities for rehearsal. The structured task group was assigned to condition 1, the pair-work group was assigned to condition 2, and the control group was assigned to condition 3. The present study addresses the following research questions:

(1) What is the effect of task rehearsal on improvements in pragmatics and speaking skills on the same (rehearsed) task?

(2) To what extent can task rehearsal influence performance on a new (unrehearsed) task?

(3) What are learners' attitudes towards CASTs and pair-work activities?

Improvements were measured by gains in pragmatic appropriateness, fluency, accuracy, and complexity.

2.7 Hypotheses

1. Given that guided planning has been found to be more effective than planning in pairs between learners, that authentic input is beneficial, and that rehearsal as a type of planning is better than no planning, it is predicted that the structured task group will outperform the pair-work group, and that the experimental groups will outperform the control group. However, since

there are different findings in previous research on the effect of task rehearsal on fluency, accuracy, and complexity, no predictions are made as to which of these aspects will improve. It is predicted that the experimental groups will improve at least on pragmatics because their attention will be drawn to this aspect during consciousness-raising activities before task planning.

2. Although there is no conclusive evidence that rehearsal improves performance on a new task of the same format (Bygate, 2001; Gass et al., 1999), Bygate (2001) stipulated that rehearsal of several tasks over a period of time might produce different results. Given that in the present study several tasks were rehearsed, it is predicted that the effect of rehearsal will transfer to new tasks. The predictions for group differences and for specific areas of improvement are the same as for research question 1.

3. Since the CASTs that are used in this study have not been utilized for teaching before, no predictions are made about learners' attitudes towards these tasks as compared to learners' attitudes towards role-plays in pairs.

The reason for investigating learners' attitudes towards classroom tasks and activities is that attitudes are a part of motivation, which is an important affective variable in SLA (Dörnyei, 2001, 2006; Gardner, 1985, 2001; MacIntyre, 2002). Investigating attitudes toward classroom activities is particularly important for adult learners who often display a wide range of learning goals, expectations, and abilities (de Saint Léger & Storch, 2009). Thus, learners' comments can inform future design of CASTs.

CHAPTER 3: METHOD

3.1 Participants

The participants in this study were students from four intact ESL classes at Michigan State University. The skills targeted in these classes were listening and speaking. All students participated in the study-related activities, which were part of the course. However, only the data from those students who consented to provide them for research were analyzed. The data from participants who were less than 18 years old were not used in accordance with policies on human research subjects. Participants who missed more than two (out of six) pre- and post-test activities were also excluded. After all exclusions, there were 17 participants in the structured task group, 16 participants in the pair-work group, and 17 participants in the control group. The control group was comprised of students from two classes rather than one because too much data had to be excluded for the reasons mentioned above. The experimental groups (classes 1 and 2) and part of the control group (class 3) participated in the study during the fall 2009 semester, and the rest of the control group (class 4) participated during the spring 2010 semester.

3.1.1 Group Comparison at the Onset of the Study

All participants were enrolled in level 2 courses at a four-level pre-academic ESL program. The students were at the same level of proficiency as determined by the listening and reading placement tests in this program – the Michigan State University English Language Test (MSUELT). The MSUELT listening section consists of two parts: short conversations and extended discourse. The extended discourse part is further divided into a lecture and a long conversation. Short conversations, a lecture, and a long conversation are all followed by a series of multiple-choice questions. The reading section consists of four passages, each followed by a

series of multiple-choice questions.[2] The groups' mean MSUELT total scores (i.e., listening and reading combined) were 59, 60, and 58 for the structured task, pair-work, and control groups respectively, as reported in Table 1 below. One-way ANOVAs indicate that there were no significant group differences on MSUELT listening scores, $F(2, 47) = 0.252$, $p = .778$, MSUELT reading scores, $F(2, 47) = 0.591$, $p = .558$, and MSUELT total scores, $F(2, 47) = 0.625$, $p = .540$.[3]

The participants' demographic information, which was self-reported on the background questionnaire, is also provided in Table 1. The background questionnaire included questions on age, gender, first language, history of language learning, and attitudes towards technology and group work (see Appendix A). The groups were fairly homogeneous in terms of the demographic factors that were surveyed. The average age for each group was 20; there were more male than female students in each group; majority of students in each group were L1 speakers of Mandarin Chinese, and the second largest group was L1 speakers of Arabic; on average, students in each group studied English from seven to eight years. The only difference between groups is that there were more L1 speakers of Arabic in the control group (five students) than in each of the experimental groups (one or two students). Participants' beliefs about the use of technology and group work for language learning were also assessed (see Appendix A for the wording of questions on the background questionnaire). There were no major differences between the groups: most of the students were comfortable with technology, preferred learning with technology rather than without, preferred studying with somebody, and thought that group

---

[2] MSUELT also has a writing section, but it is usually taken only by those students who have been conditionally admitted to the University's degree programs. The scores on the writing section are not used for placing students into ESL courses.

[3] The assumptions of normality of distribution and homogeneity of variances were tested and met.

activities were useful for language learning. However, the structured task group was slightly more comfortable with technology than the other two groups, and several participants in the pair-work group liked both studying alone and with somebody.

Table 1: Group Comparison at the Onset of the Study

	Structured Task Group (N = 17)		Pair-work Group (N = 16)		Control Group (N = 17)	
	M	SD	M	SD	M	SD
MSUELT Scores						
Listening	62	6.7	61	8	60	8.6
Reading	55	9.8	59	9.3	55	9.8
Total	59	4.2	60	4.4	58	5.4
Age	20	2.9	20	2.4	20	2.9
Gender						
Male	11		10		14	
Female	6		6		3	
L1						
Mandarin Chinese	14		12		11	
Arabic	2		1		5	
Korean	1		1			
Cantonese			1		1	
Russian			1			
Years of Studying English	7	2.4	8	3.5	8	2.6
Familiarity with Technology						
Very Comfortable	9		5		8	
Comfortable	7		7		5	
Somewhat Comfortable	1		3		4	
Prefer Learning						
With Technology	14		14		16	
Without Technology	2		1		1	
Prefer Studying						
With Somebody	10		5		13	
Alone	4		3		3	
Both Ways	2		5		1	
Depends on a situation	1		2			
Group Activities are Useful						
Yes	13		12		12	
No	1		1		2	
Depends on a situation	2		2		2	

Note. A small portion of the data is missing because one student in each group did not answer several questions.

3.2 Procedure

The researcher taught the two classes in which the experimental groups were enrolled. The same syllabus was used in these classes, and the only difference between the two groups was the experimental treatment. Two different instructors taught the classes in which the control groups were enrolled. Apart from the experimental treatment, the syllabi for these classes were similar to the ones in classes 1 and 2 as they had to conform to the curriculum established in the particular ESL program.

The experimental part of the study took place over the course of one semester (see Table 2). All students were first informed about the study, and those who agreed to participate signed a consent form. All students then filled out the background questionnaire. The participants then completed several practice activities in order to learn how to work with CASTs and how to record their interactions in pairs using *Audacity*, which is software for audio recording and editing. This training phase was included because some learners might not be literate with regard to technology use, but more importantly due to the specialized nature of the computer-assisted language learning tasks (Hubbard, 2004; Winke & Goertler, 2008).

The study included two modules: (1) apologies and (2) requests and refusals. However, for the reasons described below, class 4 (part of the control group) did not participate in the apologies module. Both modules followed the same pretest-treatment-posttest design.[4] The rationale for including these particular speech acts is that they are most often studied and are generally difficult for learners (Blum-Kulka, et al., 1989; Feak, et al., 2009; Hudson, Detmer, & Brown, 1995). Following the training phase, the students began the apology module. All groups took *computer* pre-tests in a form of CASTs and *pair* pre-tests in a form of role-plays in learner-

_____

[4] This design was employed for two reasons: to establish the equivalency of groups at the onset of the study and to measure gains between the beginning and the end of the study.

Table 2: Procedures

Timeline	Structured Task Group (Treatment = CASTs)	Pair-work Group (Treatment = PRP)	Control Group: Class 3	Control Group: Class 4
Week 2	Consent procedure; background questionnaire			
Week 3	Technology training			
Apologies				
Week 4, day 1	Computer pre-test: tasks 1A-3A			
Week 4, day 2	Pair pre-test: tasks 1A-3A	Pair pre-test: tasks 1A & 3A		
Week 5	Task 1A: treatment and pair post-test			
Week 6	Task 2A: treatment and pair post-test			
Week 7, day 1	Task 3A: treatment and pair post-test	Pair post-test: tasks 1A & 3A		
Week 7, day 2	Computer post-test: tasks 1A-3A			Consent procedure; background questionnaire
Week 8	Feedback on apologies			Technology training
Requests and Refusals				
Week 9, day 1	Computer pre-test: tasks 1R-3R		Computer pre-test: tasks 2R & 3R	
Week 9, day 2	Pair pre-test: tasks 1R-3R		Pair pre-test: tasks 1R & 3R	
Week 10	Task 1R: treatment and pair post-test			
Week 11	Task 2R: treatment and pair post-test			
Week 12, day 1	Task 3R: treatment and pair post-test		Pair post-test: tasks 1R & 3R	
Week 12, day 2	Computer post-test: tasks 1R-3R		Computer post-test: tasks 2R & 3R	
Week 13	Feedback on requests and refusals			
Week 14	Exit Questionnaire			

Note. PRP = pair role-play. A = an apology module task. R = a request and refusal module task.

learner pairs. Then the experimental groups received treatment and took a pair post-test once a

week after each of three treatment tasks, while the control group took all pair post-tests on day 1

of week 7 (see Table 2).[5] After that, all participants took a computer post-test. (The treatment

and tests are explained in details in the following sections). Since this was a semester-long

classroom-based study, the students received feedback on their performance, which was done

after the post-tests. The students then began the requests and refusals module, which followed

the same procedures as the apology module. Note that because the control group participants

could be recruited for a limited amount of time, they took slightly shorter pre- and post-tests than

did the experimental groups (i.e., they did not complete task 2 in apologies pair test, task 1 in

requests and refusals computer test, and task 2 in requests and refusals pair test). At the end of

the study, the participants in the two experimental groups took the Exit Questionnaire about their

attitudes towards treatment (see Table 2).

While the apologies module was included in the design, the data from this module were

not analyzed because a large amount of apologies data was lost. The participants for whom the

data were lost produced responses but failed to record them because they were still learning the

different technologies and getting used to them. Specifically, the data for at least two out of three

apology tasks were lost for two participants in the structured task group, five participants in the

pair-work group, and two participants in the control group. This loss of data especially affects

the pair-work group, which would have consisted of only eleven participants if the apology data

had been analyzed. Such small group size would have limited the kinds of statistical analyses

that could be performed. The inclusion of apology data would have also made group sizes

considerably uneven (15, 11, and 15 participants), which would have reduced the robustness of

---

[5] The control group completed pair post-tests all at once because the students in this group
volunteered to participate only for a limited amount of time.

ANOVA procedures (Field, 2005). Because control group participants from class 4 had limited time to participate in the study and because it had already been decided to not analyze the apologies data, class 4 did not participate in the apologies module. Since the data on apologies were not analyzed, information on this speech act is not included from this point onward.

3.3 Treatment

Experimental treatment consisted of pragmatics instruction, which was based on the models suggested by Félix-Brasdefer (2006), Judd (1999), and Olshtain and Cohen (1991). The treatment for the two experimental groups differed only on how the two groups rehearsed a task: for 15 minutes per week, participants in the structured task group rehearsed a task individually via a CAST, while the pair-work group rehearsed a task in dyads. The control group did not receive any of the treatment described below, at least not in any systematic way.[6] The experimental treatment was conducted once a week for three weeks and consisted of the following activities in this order.

1. The students first engaged in consciousness-raising activities that lasted about 20-25 minutes. To receive authentic input, students watched a video or read a dialog as a model of how a particular speech act can be performed.[7] Then in groups, students examined the transcript and answered detailed questions about the video. For example, students were instructed to find specific expressions in the transcripts that indicate politeness, to explain why the people in the video responded in a certain way, to compare these responses to the ones that would be

---

[6] The instructors who taught the students in the control group reported that they did not focus on pragmatics in their classes, although discussions of cultural issues sometimes took place.
[7] Students were provided with a dialog only in the last week of treatment because there were technical difficulties with the video; the rest of the time, students watched videos during treatment.

appropriate in their culture, to reflect on their real-life experiences with a particular speech act, etc.

2. After the consciousness-raising phase, students were presented with a situation that they would need to act out. Students first read the situation, and they were encouraged to ask any clarifying questions. Students were then given 15 minutes to rehearse the task.

a. Students in the structured task group rehearsed the assigned pragmatic situation individually via a CAST. Each activity consisted of two parts: students played the role of person X while the computer program that delivered CASTs provided responses for person Y, and then vice versa. Students were asked to rehearse both roles as many times as they could in 15 minutes.

b. The pair-work group rehearsed the given pragmatic situation through role-play in pairs. Each activity consisted of two parts: one student in each dyad played the role of person X while another student played the role of person Y, and then vice versa. The participants were asked to switch roles and rehearse both roles as many times as they were able to in 15 minutes.

3.4 Materials

3.4.1 Speech Acts and Situations

Request and refusal speech acts were used during treatment and testing. Following Brown and Levinson (1987), social situations can be categorized by three contextual factors: interlocutors' power difference, social distance, and the degree of imposition. In order to provide learners with a challenge, the degree of imposition was high for all situations in this study as it is more difficult to perform such pragmatic situations as opposed to those with low imposition (Hudson, 2001; Liu, 2006; Roever, 2006; Taguchi, 2007). For example, in Taguchi (2007), asking a classmate for a pen was considered a low-imposition request, but asking an instructor to reschedule an exam was considered a high-imposition request. In the present study, request

situations were of two kinds: with or without a difference in power. Thus, when a student talked

to an instructor, there was a power difference. When interlocutors were friends, there was no

difference in power. In all refusal situations, which consisted of refusing an instructor's

suggestion, there was a difference in power. In all situations, except in task 2 in the

consciousness-raising phase, there was no social distance between interlocutors[8] in that the

speakers knew each other.[9]

Pragmatic situations in this study were adapted from Feak et al. (2009), Hudson et al.

(1995), Sykes (2008), and Taguchi (2007). Situations for the consciousness-raising stage,

rehearsal, and testing were similar and were matched by power difference (see Table 3). For

example, for task 1, a student had to ask a friend for help with a paper (in the video during

consciousness-raising stage), with a move (during rehearsal and on the pair test), and with

finding a place to stay at in New York (on the computer test).

For each situation, the participants were provided with a somewhat broader context and

more details than is shown in Table 3 (see Appendix B for rehearsal/pair test situations and

Appendix C for computer test situations). For example, for the last situation in Table 3, the

participants were told to imagine that they have just finished talking to their academic advisor

about their coursework. They were also told that the summer course is not required, and they

were reminded to reply politely. It should be noted that students were not provided with

descriptions of situations in the consciousness-raising stage. Rather, the students watched the

---

[8] Although social distance is an important factor in pragmatics, an attempt was made to keep it
constant so that tasks would not differ on too many variables in this experimental study.
[9] Following Trosborg (1995), situations in which interlocutors knew each other at least to some
degree were categorized as -social distance, and situations in which interlocutors didn't know
each other at all were categorized as +social distance.

videos and then discussed what they thought the context was (see Appendix D for transcripts of videos).

Table 3: Pragmatic Situations for Requests and Refusals

Study Phase	Task	Situation
Consciousness raising	Task 1: Request, - P	Ask your friend for help with editing a paper for a class.
	Task 2: Request, + P	Ask an instructor for permission to enroll in her class because the class is full.
	Task 3: Refusal, + P	Refuse your instructor's strong suggestion to present your paper at a conference.
Pair tests; rehearsal	Task 1: Request, - P	Ask your friend to help you move this weekend.
	Task 2: Request, + P	Ask your instructor if you can change the topic of your paper even though the paper is due soon.
	Task 3: Refusal, + P	Refuse your academic advisor's strong suggestion to apply for an internship in the summer.
Computer tests	Task 1: Request, - P	Ask your friend if you can stay at his parents' place in New York.
	Task 2: Request, + P	Ask your instructor for help on homework knowing that he had to cancel office hours this week.
	Task 3: Refusal, + P	Refuse your academic advisor's strong suggestion to take a summer course.

Note. P = power difference.

3.4.2 Videos

Videos of interactions on speech acts were recorded by native speakers of English. Native speakers were asked to perform the pragmatic situations provided to them in the way they would handle these situations in real life. That is, native speaker role-plays were unscripted in order to make them as naturalistic as possible. These videos were used for consciousness raising, as well as for constructing prompts in CASTs.

3.4.3 CASTs

CASTs were created and delivered via a computer program called *Conversations,* which was developed at the Center for Language Education and Research (CLEAR) at Michigan State University. First, the content for CASTs was created by recording a series of utterances, or

prompts, within the *Conversations* program. Each CAST included three to five prompts. These prompts were first video-recorded by the researcher based on the utterances from the role-plays performed by native speakers of English (following Sykes, 2008). The resulting CASTs were then pilot-tested with both native speakers and non-native speakers of English. Based on the responses, the prompts were revised. The final prompts were video-recorded by native speakers of English. This extensive piloting of materials was necessary to determine which prompts to record so that the prompts (or rejoinders) following the participants' responses resembled a logical or natural continuation of discourse. That is, because the rejoinders were not previewed but rather appeared only after a participant produced a response, the rejoinders which are frequently produced by a large number of people had to be determined via pilot-testing. The final prompts and the corresponding situations used for rehearsal are provided in Appendix B, and those prompts and situations used on computer tests are provided in Appendix C.

The interface of CASTs as they were presented to the participants is shown in Figures 1 and 2 below. As can be seen from both figures, the situation that the students needed to respond to was always shown at the top of the screen. Once the students were ready to begin a CAST, they would click the "Real Time" button and get the first prompt. The screen shot of what the students saw on the computer screen as they were presented with a prompt is provided in Figure 1. Figure 2 displays what students saw right after they heard the prompt and as they were recording their response. Participants recorded their response to each prompt, and they were given unlimited time to do so. After responding, students had to click the "Stop Recording" button to proceed to the next prompt. The prompts were always presented in the same order. Once students clicked the "Stop Recording" button after the last prompt, a CAST would end, and students' responses would automatically appear in the instructor's module. Students were asked

33
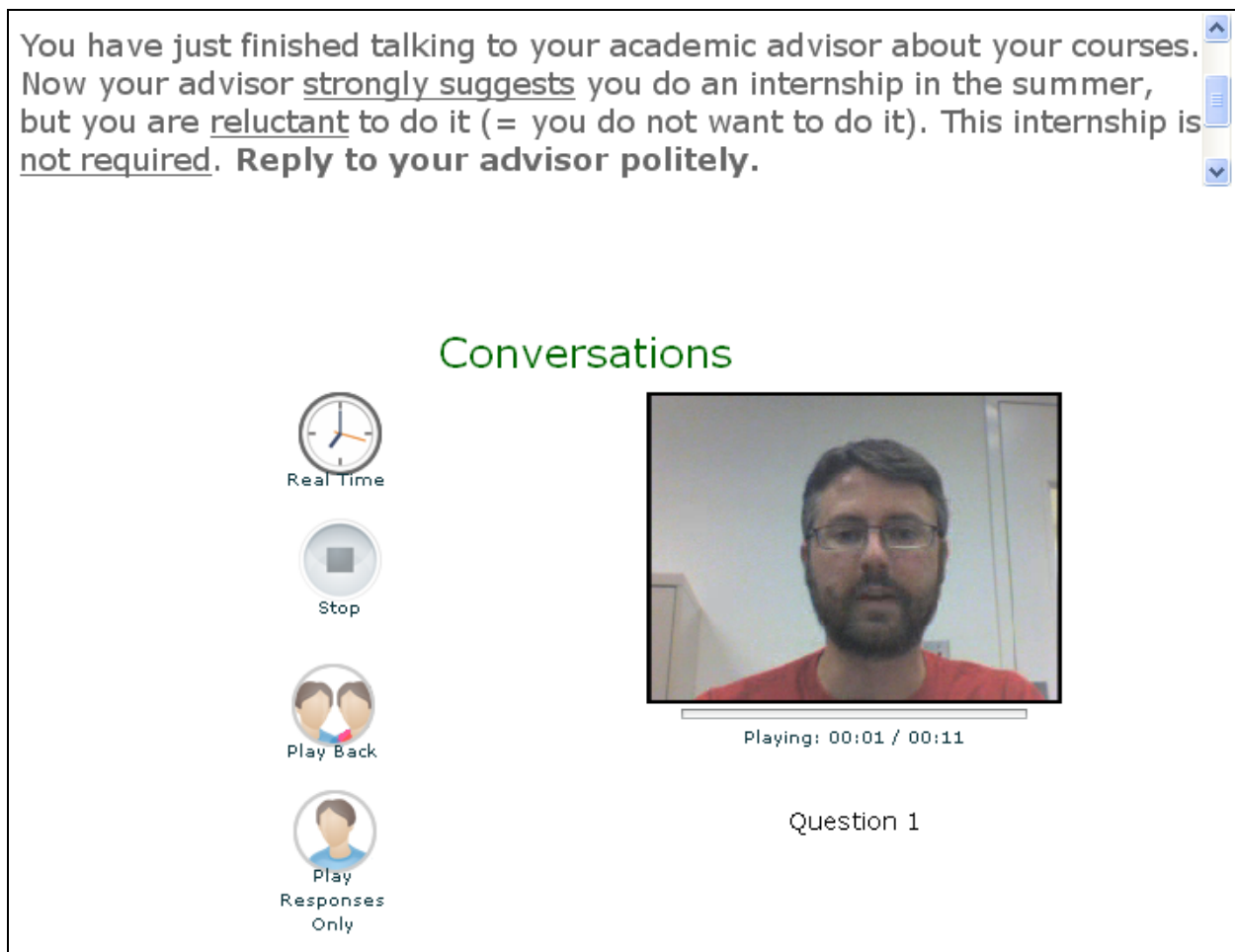
Figure 1. A Screenshot of a CAST When a Prompt is Presented

For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation. The students heard "Also, there is one other thing I wanted to talk to you about. It looks like there's a possibility for an internship in Washington DC. I think you'd be a great candidate. Would you be interested?"

to finish the whole task from beginning to end; that is, they were asked to not click the "Stop" button in the middle of a CAST or to preview the rejoinders before responding to them.[10] Once participants completed a CAST, they were asked to click either the "Play Back" or the "Play Responses Only" buttons to make sure that their responses were recorded.

---

[10] Because during rehearsal participants had to complete each CAST several times, they essentially previewed the rejoinders during the first enactment of the CAST.

During treatment, participants used CASTs to rehearse both roles in a given situation. For example, when rehearsing task 3, participants in the structured task group first acted as a student, then as an advisor, then again as a student, and so on. However, when taking pre- and post-tests, all participants only acted out one role one time (either making a request or refusing a suggestion).

You have just finished talking to your academic advisor about your courses. Now your advisor strongly suggests you do an internship in the summer, but you are reluctant to do it (= you do not want to do it). This internship is not required. **Reply to your advisor politely.**

Conversations

Real Time

Stop

Play Back

Play Responses Only

Recording time: 00:01
Stop Recording
Question 1

Figure 2. A Screenshot of a CAST When a Response to a Prompt can be Recorded

3.5 Instruments

3.5.1 Pair Tests

All students took pre- and post-tests in a form of role-plays in pairs. The experimental groups also rehearsed these tasks (either in pairs or on the computer) between the pre- and the

post-tests. Participants were always paired with classmates; that is, they worked with other participants within their group (the structured task group, the pair-work group, or the control group). To the degree possible, learners in each pair were from different L1 backgrounds. Participants were paired with different partners over the course of the semester, and those in the pair-work group had a different partner during rehearsal and during the testing phase. This design was chosen to increase the validity of role-plays as a means of measuring learners' pragmatic and speaking skills because performance depends on interlocutors (Brooks, 2009; Davis, 2009; Lazaraton & Davis, 2008; McNamara, 1997; Swain, 2001). Additionally, this design is ecologically valid since in a classroom environment students work with a variety of other students during the course of the semester. The results from pair tests were used to evaluate the effect of rehearsal on the performance of the same (rehearsed) tasks.

3.5.2 Computer Tests

In addition to pair tests, all participants also took pre- and post-tests individually on the computer via CASTs. These CASTs were slightly shorter than the ones used during treatment; they consisted of a sequence of three or four prompts. A combination of DCTs and role-plays was used for testing pragmatics because both instruments have drawbacks (Gass & Houck, 1999; Félix-Brasdefer, 2010; Sykes, 2008). (Recall that CASTs are like DCTs, except rejoinders are not shown beforehand.) Role-plays, as opposed to DCTs, elicit the data that more closely resembles natural interactions (Gass & Houck), but role-plays are also harder to score due to the interlocutor influence (see Roever, 2011, for a review). Note that the situations on the pair tests and computer tests were similar, but not the same. Since participants did not rehearse the tasks that appeared on the computer tests, their performance on these tests was used to evaluate the

effect of rehearsal on the performance of new tasks (i.e., whether task rehearsal facilitates pragmatic and oral development in general).

3.5.3 Exit Questionnaire

The Exit Questionnaire on attitudes towards treatment was adapted from questionnaires used by Kost (2004), Kraemer (2008), and Sykes (2008) in their classroom research studies, which involved speaking and/or pragmatics instruction using technology. The questions on this instrument were about role-play activities in dyads and CASTs (see Appendix E). To ensure that students felt comfortable providing feedback on classroom activities, respondents could remain anonymous on this questionnaire if they preferred.[11]

3.6 Measures

Several measures were used to examine the participants' pragmatic and oral development.

3.6.1 Pragmatic Appropriateness

The researcher (an ESL instructor who is a NNS of English) and an outside rater (an ESL instructor who is a NS of English) rated the responses in terms of pragmatic appropriateness. The pragmatic appropriateness rating scale was adapted from Hudson et al. (1995) and Taguchi (2007) and is shown in Table 4.

Following Taguchi (2007), "Appropriateness was defined as the ability to perform language functions appropriately in a social context" (p. 120). According to Hudson et al. (1995), appropriateness subsumes the amount of information provided, formality, directness, and politeness. In this study, components of pragmatic appropriateness were not explicitly stated in

---

[11] As a consequence of this, anonymous responses of several participants on the Exit Questionnaire could not be linked to their responses on the background questionnaire. Only several responses were anonymous because other participants did not opt for the anonymous option. On the other hand, the advantage of this decision was that participants felt more comfortable providing honest opinions, which can help improve the design of CASTs for future use.

the rating scale.[12] Rather, the researcher and an outside rater interpreted on their own what is considered pragmatically appropriate. However, the researcher and an outside rater also provided comments for each response explaining why they gave a particular rating (following Tateyama, 2001). These comments indicate that both raters did consider the four components of pragmatic appropriateness stated above: the amount of information provided, formality, directness, and politeness.

Table 4: Pragmatic Appropriateness Rating Scale

Ratings	Descriptors
1	Extremely inappropriate
2	Inappropriate
3	Somewhat inappropriate
4	Somewhat appropriate
5	Appropriate
6	Completely appropriate

Just like in Taguchi (2007) and Tateyama (2001), the rating scale descriptors were very general. It was decided not to impose specific guidelines on the raters in terms of factors that contribute to pragmatic appropriateness because the goal was to explore what affects raters' judgments. Two examples for each point on the rating scale, which were taken from the data in this study, are provided below. There was complete rater agreement on these examples; that is, both raters gave the same ratings. The general principles that the raters operated on were the following. To receive the rating of 6 (completely appropriate), the students really had to "shine" and be exceptionally considerate and polite. For the rating of 5 (appropriate), the students had to be simply appropriate. The rating of 4 (somewhat appropriate) was assigned to those samples which were generally appropriate, but needed slight improvement on one (or rarely two) aspects;

---

[12] This approach was taken because, as Hudson et al. (1995) acknowledge in the training guide for raters, it can be very difficult to separate the concepts of formality, directness, and politeness when rating them.

for example, when the students needed to be a bit more indirect, formal, or appreciative. The ratings of 3, 2, and 1(somewhat inappropriate, inappropriate, and extremely inappropriate respectively) were assigned when students were inappropriate, but to varying degrees. When examining the sample responses below in Examples 1 through 12, the readers should keep in mind that participants' intonation (e.g., understanding, appreciative, whiny, disappointed) was often a factor and cannot be adequately represented here. If intonation made a difference in the rating of a particular response, according to raters' comments, a note in the transcript is provided.

In Example 1, the task is to ask a friend for a favor: to stay at his parents' place.

Example 1. Sample response 1 with a pragmatic appropriateness rating of 6

(completely appropriate): computer test, participant ID 80

Prompt: Hey, what's up?

Participant: Yeah, I'm good. How how about you? Um yeah, do you have time, do you have time? Uh could I ask a favor if you don't mind? Uh you know what? I got a I got an important interview in New York in this weekend. But uh I don't have place I don't have any place to stay there. So I heard that your parents live in New York city. So uh I wanna ask you about that uh if I can live your parents' house, just couple days.

Prompt: Uh, I mean, congratulations. It's great that you got that interview and all. And I'd love to help, but I'd have to ask my parents first because their apartment's kinda small.

Participant: Yes, yes, I totally understand that. Uh if you can uh I'd really appreciate that if you if you can ask your parents uh if I can stay your parents' house.

Prompt: Yeah, I don't know if they are gonna be home. They kinda go out of town

39

a lot, but I'll call them tonight and let you know.

Participant: Okay, thank you very much. When you get a when when you get a

response, just let me know.

Prompt: Yeah, sure, no problem. I'll let you know what they say.

Participant: Thank you very much.

In Example 2, the task is to ask an instructor for help on homework.

Example 2. Sample response 2 with a pragmatic appropriateness rating of 6

(completely appropriate): computer test, participant ID 131

Prompt: Okay, what's your question?

Participant: [Sigh] uh actually I I have the question and and I it's emergency to ask

you about it it's 'cause it's very very very difficult for me. I asked I asked

the friends around me but they ha- all have no answer. Uh so I just

wondering that if if you can spend five or ten minutes to explain this

question. I I know you are very busy this week 'cause you have to cancel

the office hour yeah this is why I come here to ask.

Prompt: Oh, I'm I'm really sorry, thi- this week is very bad for me. I'm extremely

busy. And in fact that's why I had to cancel my office hours. What I'd

recommend is that you talk to the tutor. He is very good at explaining

these things.

Participant: Oh I have tried it be he's cannot explain it clearly. Um yeah I I know you

are you are very busy but I I just need five minutes for you to uh explain it

'cause I have I have collect all the information around maybe and it can

probably (can save?) your time.

40

Prompt: I'm afraid I'm really pressed for time right now and in fact I've gotta go teach in a few minutes. And and I truly apologize but this week as I said is really busy. I'm I'm leaving town, so I'm not gonna be able to meet with you in person. But if you'd like to email me, I'd be more than happy to uh to respond to you. And also you might wanna check the PowerPoints that we used in class. You could also use the software in the language lab.

Participant: Oh that's great! You you can uh you can answer me from the email? Oh I I cou- I could email you, uh yeah I can explain the question on on the e- on the email box. Thank you so much! You are so nice.

Prompt: Yeah, and we can talk during my office hours next week.

Participant: Okay, that's great, no problem. [Excellent intonation: shows appreciation and understanding.] And there maybe I maybe I just to email you 'cause the s-s the time is too limit, I have to hand in the homework.

In Example 3, the task is to decline an instructor's suggestion.

Example 3. Sample response 1 with a pragmatic appropriateness rating of 5
(appropriate): pair test, participant ID 40

Partner: Come in.

Participant: Hi, Professor.

Partner: Hi. Uh the another one is I would like to told you that you have a chance to do the internship in this summer in Washington DC. Uh how about (feeling?) in uh to do this?

Participant: Oh. I am very appreciate it you can told me that. But I have another plan this summer. But because I want to make more money for myself. So uh I

41

am not sure.

Partner: Oh uh I can understand you but I think uh this ti- uh this chance is very

good for you. I think it's a goo- uh right chance because you can improve

uh some experience in your life uh and I think it's the better chance.

Participant: Yeah. I know. So but uh if I don't have money, I can't graduate from my

university. So I think it's very good, very important things. So

Partner: Uh yes uh I can understand you. I k- uh I'm very (fond?) to help you on

competition uh but thi- this uh this decision due in the two weeks you can

tell me the quickly uh I'm very (fond?) to help you this time because I

think it's really really uh improve you you major or the something.

Participant: Oh. Okay. Uh thanks a lot. If I decide to do this, I will let me let you

know soon. So um I am very appreciate you can told me that.

Partner: Okay I really uh have more happy that uh you make the right decision and

if you have some problems you can tell me.

Participant: Okay. Thanks a lot.

Partner: That's okay.

Participant: Yeah. Thanks a lot.

Partner: You are welcome.

In Example 4, the task is to ask a friend for a favor: to stay at his parents' place.

Example 4. Sample response 2 with a pragmatic appropriateness rating of 5

(appropriate): computer test, participant ID 81

Prompt: Hey, what's up?

Participant: Uh hey uh I have a interview in New York uh next week. So I I remember

your parents live here so I wondering if I can stay in your parents' home for a few days, please.

Prompt: Uh, I mean, congratulations. It's great that you got that interview and all. And I'd love to help, but I'd have to ask my parents first because their apartment's kinda small.

Participant: Yeah um of course you should ask them first because you know I go to New York and the hotel is too expensive. And I just (will?) stay a few days so I wonder if you can help me.

Prompt: Yeah, I don't know if they are gonna be home. They kinda go out of town a lot, but I'll call them tonight and let you know.

Participant: Okay, if you do that, thank you a lot.

Prompt: Yeah, sure, no problem. I'll let you know what they say.

Participant: Okay, thank you.

In Example 5, the task is again to ask a friend for a favor: to stay at his parents' place.

Example 5. Sample response 1 with a pragmatic appropriateness rating of 4 (somewhat appropriate): computer test, participant ID 141

Prompt: Hey, what's up?

Participant: Uh I'm just I wonder if I can stay at your parents' place in New York. You know I have a interview in New York so I have to find a place to live.

Prompt: Uh, I mean, congratulations. It's great that you got that interview and all. And I'd love to help, but I'd have to ask my parents first because their apartment's kinda small.

Participant: Yeah, I know, so okay. [Intonation expresses disappointment rather than

43

understanding]

Prompt: Yeah, I don't know if they are gonna be home. They kinda go out of town a lot, but I'll call them tonight and let you know.

Participant: Thank you, thank you very much.

Prompt: Yeah, sure, no problem. I'll let you know what they say.

In Example 6, the task is to ask an instructor for permission to change the topic of the paper. Note that whenever names are provided, they are pseudonyms.

Example 6. Sample response 2 with a pragmatic appropriateness rating of 4 (somewhat appropriate): pair test, participant ID 161

Partner: Hey Jason.

Participant: Hey my teacher.

Partner: Uh do you have any problem?

Participant: Yeah I want to change change my top topic of uh topic.

Partner: Oh you want to change your topic? Why?

Participant: Uhu. Uh because I think uh the original topic is not fit on me and I think the new topic is more interesting and you would like the new topic I think.

Partner: But

Participant: Uh do you want do you want

Partner: Don't you think uh it too late? Uh the paper is due soon.

Participant: Yeah I know but I can I can I can finish it now if I change the topic.

Partner: But I don't think uh you will do it well it's uh you don't have enough time to do it.

Participant: I I can do it it trust me. [Due to intonation, sounds whiny.] Uh yeah I I

like uh new topic and I can I can do it as well as um uh even if

Partner: Uh why don't you change change your topic earlier? Um now it's too late. Uh maybe next time you can chan- change the topic and do a new new one.

Participant: Uh you know uh even if I do that original topic I I can't do it very well I I I don't interesting in that and I think that topic is not fit on me. Uh if you want

Partner: That's a challenge for you. You can you can uh you can do your best.

Participant: Okay.

Partner: I trust you.

Participant: Okay okay. Thank you.

Partner: Okay. See you.

Participant: Okay, see you.

In Example 7, the task is to ask a friend for a favor: to stay at his parents' place.

Example 7. Sample response 1 with a pragmatic appropriateness rating of 3 (somewhat inappropriate): computer test, participant ID 161

Prompt: Hey, what's up?

Participant: Hey, I remember you told me your parents uh live in New York. You know what? I have a interview in New York and this weekend and this is very important for me. And uh could you uh told uh told your tell your parents and I want to stay your parents' home uh just one just for a couple of day.

Prompt: Uh, I mean, congratulations. It's great that you got that interview and all.

45

And I'd love to help, but I'd have to ask my parents first because their apartment's kinda small.

Participant: Okay, uh okay.

Prompt: Yeah, I don't know if they are gonna be home. They kinda go out of town a lot, but I'll call them tonight and let you know.

Participant: Yeah, thank you, thank you very much. Uh I will waiting for your answer.

Prompt: Yeah, sure, no problem. I'll let you know what they say.

Participant: Okay, thank you.

In Example 8, the task is to decline an instructor's suggestion.

Example 8. Sample response 2 with a pragmatic appropriateness rating of 3 (somewhat inappropriate): pair test, participant ID 141

Partner: Hi Julie.

Participant: Hi Kathy.

Partner: I think you do very well about your courses. Now I want to

Participant: Thank you.

Partner: Now I want to talk to you about an internship this summer in DC.

Participant: Okay.

Partner: Uh do you want to take this one?

Participant: Uh it is required to student must to attend?

Partner: No, it's not required.

Participant: Oh, so I think I prefer not.

Partner: But I think uh it's very important uh because if you take an internship you can got uh get a lot of experience. It's good for your uh work when you

graduate from the college.

Participant: Maybe maybe but I think it is not required so I don't have to do it and I have little time to do it because I have busy courses in summer.

Partner: Oh you have a courses?

Participant: Yes in summer so I'm not sure I have enough time to do this internship so

Partner: Ok uh if you have uh time you can do it another time.

Participant: Yes okay, bye.

In Example 9, the task is to ask an instructor for help on homework.

Example 9. Sample response 1 with a pragmatic appropriateness rating of 2 (inappropriate): computer test, participant ID 61

Prompt: Okay, what's your question?

Participant: Oh um instructor you know uh this um this week's homework is so difficult for me and I know you are very busy um so you canceled the o- uh office hours (this morning?). But I really need your help for this homework. Uh can you um make schedule for me or um something else to help me to do jo- do the homework? Thank you.

Prompt: Oh, I'm I'm really sorry, thi- this week is very bad for me. I'm extremely busy. And in fact that's why I had to cancel my office hours. What I'd recommend is that you talk to the tutor. He is very good at explaining these things.

Participant: Oh tutor? Oh I I know that but I I want however I want to you to help me. I think you your explain is b- better for uh better for me to understand that. Um I just need a little time to do that. Uh I think it's doesn't matter

47

Prompt: for you to help me. That's very important for me and I really want you can help me. So please make (such a?) schedule for me. Just half hour, it's ok.

Prompt: I'm afraid I'm really pressed for time right now and in fact I've gotta go teach in a few minutes. And and I truly apologize but this week as I said is really busy. I'm I'm leaving town, so I'm not gonna be able to meet with you in person. But if you'd like to email me, I'd be more than happy to uh to respond to you. And also you might wanna check the PowerPoints that we used in class. You could also use the software in the language lab.

Participant: Okay, thank you so much. I think I will email you. And I hope you can have a time to uh read my homework and give s- uh give me some advantages. Uh thank you so much.

Prompt: Uh yeah, and we can talk during my office hours next week.

Participant: Okay, thank you so much. So next week I will come to your office hours. And I hope you can have enough time to read my email and give me the right advantages.

In Example 10, the task is to ask a friend for help with the move to another place.

Example 10. Sample response 2 with a pragmatic appropriateness rating of 2 (inappropriate): pair test, participant ID 183

Partner: Hey man, how's it going?

Participant: Not much. This weekend I I have to m- uh move some stuff to another hall. And so I need your help.

Partner: Oh I'm really sorry in this weekend I have something else to do.

Participant: Oh but I have lot of uh lot of stuff need to need to move. Uh I I hope you

can help me.

Partner: Oh maybe I can find somebody else to help you.

Participant: Okay.

Partner: Okay.

In Example 11, the task is to ask a friend for help with the move to another place. Note that the rating of 1 was given by both raters in only three cases.[13]

Example 11. Sample response 1 with a pragmatic appropriateness rating of 1 (extremely inappropriate): pair test, participant ID 23

Participant: Hi Ben, how are you?

Partner: Oh hi Dave, I'm fine. How about you?

Participant: Well, not much. Uh I will move to You Nyork uh next week. Can you help me?

Partner: Oh, really? It's good idea. What what can I do for you?

Participant: Please I please I want to help me in the moving to the my new apartment.

Partner: Oh, moving? Uh, what's time? Next week?

Participant: Next week in Friday.

Partner: Oh my God. But I have I have eh another plan. I I promise my wife go to uh California.

Participant: Please, I need your help. You are my best friend.

Partner: But it's difficult. You know it's my wife. I promise my wife. I I I'm afraid him.

---

[13] Note that qualitatively, responses with a rating of 1 look like outliers. Thus, an additional statistical analysis was conducted in which responses with a rating of 1 were excluded. This analysis is reported on in chapter 4.

Participant: Promise her to another day.

Partner: Um okay.

Participant: Please Ben try your best.

Partner: But I need I need help to call her. I I can't do that by myself.

Participant: That's my cellphone. Please call her now.

Partner: Okay. Hello. Ah oh okay okay. See you. Oh my wife uh go to another place.

Participant: Yeah that's look great.

Partner: And with the new boyfriend. Oh my God.

Participant: So you can help me now.

Partner: Yeah, of course.

Participant: Okay, that's great. Thank you.

Partner: So see you guys.

Participant: See you.

In Example 12, the task is again to ask a friend for help with the move to another place.

Example 12. Sample response 2 with a pragmatic appropriateness rating of 1

(extremely inappropriate): pair test, participant ID 171

Participant: Hi Dave. How's it going?

Partner: Hi. Good and how are you?

Participant: Good. Uh I was wondering do you have a time to help me move my move my stuff this week this weekend?

Partner: Um maybe I could I'm sorry maybe I could not because I have some other plans. I promise other peoples to take take uh to take them to Detroit so I

have no time to help you to move. Because I I you know I so they they

aks me they ask me could I could I could I could I took them.

Participant: So when you come back?

Partner: Um, maybe on Monday.

Participant: Oh, that's really bad. But do you have some friend can help me?

Partner: Um, maybe not? Let me think.

Participant: You don't have friend?

Partner: Let me think. Uh maybe yes um maybe Luke have free time. Maybe Luke

will have Luke has free time free time.

Participant: But but his car is too small. He can't. I have I have a huge bag and huge

TV. His car can't do that.

Partner: Maybe Tom can. You can you can ask him. I think you aks him.

Participant: Okay, I just ask you.

Partner: He always can help others. He always help others. He is a kind of man.

He he is a kind of man.

Participant: I just ask you. You never help others. So see you.

Partner: See you.

Pragmatic appropriateness scores were calculated as the mean of the two ratings.

Following Taguchi (2007), raters discussed the responses for which the ratings differed by more

than one point and adjusted their ratings accordingly. For one case when agreement could not be

reached, the judgment of a third rater (an ESL instructor who is a NS of English) was used. After

resolving disagreements, average interrater reliability for all tasks combined was .87 (Pearson r,

two-tailed). The interrater reliability for each task is provided in Table 5.

51

Table 5: Interrater Reliability of Pragmatic Appropriateness Ratings

	Pre-Test	Post-Test
Pair Test		
Task 1	.90	.89
Task 2	.83	.87
Task 3	.82	.88
Computer Test		
Task 1	.81	.83
Task 2	.93	.92
Task 3	.89	.87

3.6.2 Oral Development

Analytic measures of fluency, complexity, and accuracy are generally used in planning studies on oral development, and they were also employed in this study.

Fluency is a complex construct to measure. Skehan (2003, 2009) and Tavakoli and Skehan (2005) believe that there may be three distinct aspects of fluency: (1) breakdown fluency, which involves silences; (2) speed fluency, which is calculated by different speed and time-related measures; and (3) repair fluency, which involves the examination of reformulations, replacements, false starts, and repetitions of words and phrases. Segalowitz (2010) also points out that the situation is complex in terms of objective measures of fluency: "despite several decades of work, researchers have not discovered universally applicable, objective measures of oral fluency" (p. 39). However, he also states that speech rate and number and length of silent pauses have the strongest correlations with oral proficiency. Such findings were obtained by Iwashita, Brown, McNamara, and O'Hagan (2008), Rossiter (2009), and Thomson and Isaacs (2010), who compared different objective measures of fluency with human fluency ratings. In fact, speech rate might be either the best measure of fluency or as good as others. In Iwashita et al. (2008), proficiency level showed the strongest correlation with speech rate, followed by silent pause rate and total pause time. Bhat, Hasegawa-Johnson, and Sproat (2010) found that speech

rate correlated with the human-rated fluency scores as much as the measures based on silent pauses. Following Segalowitz's (2010) review of fluency measures, speech rate and the number and length of silent pauses will be the measures used in this study.

The majority of planning studies have used two measures of speech rate: syllables per minute (or per second) and pruned syllables per minute (or per second) (e.g., Mehnert, 1998; Ortega, 1999; Sangarun, 2005; Yuan & Ellis, 2003). In some studies, such as Guará-Tavares (2008) and Mochuzuki and Ortega (2008), words rather than syllables were counted. In the measure of pruned speech rate, repetitions, reformulations, and false starts are excluded from the syllable or word count (Mehnert; Yuan & Ellis). In this study, speech rate was measured as syllables per minute and as pruned syllables per minute. Pruned speech rate is viewed as a "measure of meaningful productivity" of language (Yuan & Ellis, p. 13). Following Derwing, Munro, Thomson, and Rossiter (2009), Lennon (1990), Mochuzuki and Ortega, and Rossiter (2009), repetitions, reformulations, and false starts, as well as non-lexical fillers (such as *um* or *uh*), asides, and utterances in L1 were excluded from pruned syllable count. However, repetitions for rhetorical effect (e.g., Ortega) were kept. Syllables were counted using the website http://www.addedbytes.com/code/readability-score/. There was 100% agreement between the manual count of syllables in five speech samples and the syllable count provided by this tool.

The number of pauses and total length of pauses were also calculated, following many planning studies (e.g., Foster, 1996; Foster & Skehan, 1996; Mehnert, 1998; Skehan & Foster, 1997; Tavakoli & Skehan, 2005; Wendel, 1997). To take into account varying lengths of responses, number of pauses per minute and percentage of pause time were the exact measures employed (e.g., Iwashita et al., 2008; Mehnert). Different researchers used different thresholds for the minimum duration of a pause, from 200 to 1000 milliseconds. Following Foster and

Skehan (1996, 1999), Iwashita et al., Mehnert, Skehan and Foster (1997), Sundqvist (2009), and

Weissheimer (2007), pauses of one second or longer were analyzed. Phonetic analysis software

Praat was used to determine the length of pauses.

In planning studies, many researchers examine both syntactic and vocabulary complexity.

The most common measure of syntactic complexity is the ratio of clauses to either T-units, c-

units, or analysis of speech (AS) units (R. Ellis, 2009). In this study, the ratio of clauses to AS-

units is used, following Michel, Kuiken, and Vedder (2007) and Tavakoli and Skehan (2005).

The AS-unit was introduced by Foster, Tonkyn, and Wigglesworth (2000). AS-units are similar

to T-units or c-units, with several modifications, which allow for the coding of highly interactive

discourse in a principled way.[14] All guidelines provided by Foster et al. (2000) for dividing the

data into AS-units were followed. Additionally, Foster et al. state that researchers using

interactional data often felt the need to exclude some parts of the data from analysis, and thus

Foster et al. provide three levels of application of the AS-unit to make the data exclusion process

more principled. Level three was used in this study, which Foster et al. consider necessary when

comparison of participants on quantitative measures is performed. That is, the following

elements were excluded: (a) one-word minor utterances and one-word elliptical utterances (e.g.,

"Yes;" "No;" "Uhuh;" "Sunday" (as an answer to the question "When do you want to move?"));

(b) echo responses which are verbatim; (c) one- or two-word greetings and closures.

Clauses were identified by closely following the guidelines provided in Foster et al.

(2000). However, unlike Foster et al., Bardovi-Harlig and Bofman (1989), Hunt (1970), Polio

(1997), and Wolfe-Quintero, Inagaki, and Kim (1998) state that a clause must contain a finite

---

[14] See Norris and Ortega (2009) for a discussion of appropriate contexts for using T-units, c-
units, and AS-units.

verb, while non-finite verbs cannot constitute a clause. In the current study, non-finite verbs also did not constitute clauses.

For vocabulary complexity, most studies have used type-token ratio. However, some researchers are beginning to use Guiraud's Index when the length of responses varies widely (e.g., Gilabert, 2007; Michel et al., 2007; Sauro & Smith, 2010). Guiraud's Index is defined as the number of types divided by the square root of the number of tokens. That is, Guiraud's Index accounts for the difference in the length of responses, while type-token ratio does not (Dewaele & Pavlenko, 2003; Vermeer, 2000). The higher the Guiraud's Index, the higher is the lexical complexity (or diversity) of a given response. Following Iwashita et al. (2008), the number of types and tokens was calculated using the web program *VocabProfile* (Cobb, 2002).

To examine changes in accuracy, the most frequent measure, ratio of error-free clauses, was used (e.g., Elder & Iwashita, 2005; Sangarun, 2005; Skehan & Foster, 2005; Tavakoli & Skehan, 2005). Other measures that have been previously used are the number of errors per 100 words (see R. Ellis, 2009) and changes in accuracy on specific grammatical issues, such as verbs *ser* and *estar* in Spanish (Gass et al., 1999) or relative clauses in English (Mochizuki & Ortega, 2008).

3.7 Analysis

In order to examine both main effects and interactions, quantitative analyses were conducted using one mixed-design ANOVA for each dependent variable (i.e., speech rate, pruned speech rate, number of pauses per minute, percentage of pause time, ratio of clauses to AS-units, Guiraud's Index, and ratio of error-free clauses), with time (pre-test, post-test) as a within-subjects independent variable and treatment group (structured task, pair-work, control) as a between-subjects independent variable. That is, a mixed-design ANOVA procedure produced

the following three pieces of information for each dependent variable: (a) whether the participants improved regardless of the group; (b) whether the groups differed on gains between the pre-test and the post-test; and (c) whether the groups differed on their combined pre-test and post-test performance. It is important to know whether there was an overall improvement on each dependent measure because before comparing the groups on gains, one needs to know if there were any significant gains. The comparison of groups on gains directly answers research questions 1 and 2. The comparison of groups on their combined pre-test and post-test performance is not directly related to research questions, but it can help explain the results. To control for Type I error, which may be caused by the large number of tests conducted on the same data set, the Bonferroni correction was applied (e.g., Keselman et al., 1998). Thus, the alpha level of .05 was divided by eight (the number of mixed-design ANOVAs), and the alpha level used to test the significance of mixed-design ANOVAs in this study was .006.

Mixed-design ANOVAs were then followed up with ANCOVAs for those variables on which the groups differed on the pre-tests. To examine whether there were any group differences before the treatment, one-way ANOVAs were conducted on the pre-test scores for each of the eight dependent variables. (The Bonferroni correction was applied to the one-way ANOVAs conducted on the pre-test scores; thus the alpha level of .006 was used to test the significance of one-way ANOVAs). In pretest-posttest designs, when pre-test scores differ between groups due to pre-existing group differences, ANCOVAs (with pre-test scores as a covariate) are commonly used to eliminate systematic bias (e.g., Hatch & Lazaraton, 1991; Hinkle, Wiersma, & Jurs, 2003; Mackey & Gass, 2005), but some caution against such use of an ANCOVA (e.g., Field, 2009; Pedhazur, 1997). Due to this disagreement among researchers, both mixed-design ANOVA and ANCOVA tests were used.

The analyses on the average scores for the three rehearsed tasks and the average scores for the three new tasks were performed. That is, following Wigglesworth (1997) and Takimoto (2009), scores for similar tasks were collapsed: the scores were averaged for the three rehearsed tasks (one set of scores) and for the three new tasks (another set of scores). Combination of tasks by type is common in language testing (e.g., Brown, Iwashita, & McNamara, 2005). Ratings on different speech acts have been also summed up previously (e.g., Taguchi, 2007). The reason for collapsing similar tasks is to make the assessment of changes more reliable rather than idiosyncratic. The scores were collapsed because the graphic representation of descriptive statistics revealed similar trends across the rehearsed tasks and across the new tasks. Another reason for averaging the scores of the rehearsed tasks and those of the new tasks is that some data were missing. Specifically, one participant from each of the experimental groups did not complete one of the tasks and the control group performed two out of three rehearsed tasks and two out of three new tasks. The group sizes for rehearsed tasks were 17, 16, and 17 for the structured task, pair-work, and control groups respectively. However, the groups sizes for new tasks were 17, 15, and 16 (the structured task, pair-work, and control groups respectively) because one participant in the pair-work group and one participant in the control group performed but failed to record their responses on the computer tests.

The assumptions for ANOVA tests were checked. ANOVA is reliable when (1) observations are independent, (2) the dependent variable is measured on at least an interval scale, (3) the data within each group are from a normally distributed population, and (4) variances in each group are similar (Field, 2005). The first two assumptions were met: observations were independent in that no participant was associated with more than one experimental group, and the dependent variables were not dichotomous or categorical.

The assumption of normal distribution was not met for some dependent variables. The data for both rehearsed and new tasks were not normally distributed for number of pauses per minute and percent of pausing time, and various transformations did not make the data more normally distributed. Thus the results on pausing need to be interpreted with caution. There were also minor violations of normality for new, but not for rehearsed, tasks on syllables per minute and ratio of clauses per AS unit. That is, the Shapiro-Wilk test was significant at the alpha level of .05 only for one of the three groups and only on either the pre-test or the post-test. The logarithmic transformation of the ratio of clauses per AS unit made the data more normally distributed for this variable as determined by the Shapiro-Wilk test.

The homogeneity of variance assumption has been violated (i.e., $p < .05$ on the Levene's test of equality of variance) for speech rate and pruned speech rate on both rehearsed and new tasks. Examination of descriptive statistics revealed that standard deviations for the control group were higher than those of the experimental groups. That is, there was more variability within the control group than within the other two groups on speech rate. Thus, the results on speech rate should be interpreted with caution, especially when comparisons are made with the control group.

Even though the assumptions of normal distribution and homogeneity of variances were not always met, ANOVAs were conducted because (1) these assumptions were violated only for some dependent variables; (2) ANOVA is robust with respect to non-normal distribution (i.e., violating this assumption may not necessarily be problematic); and (3) when group sizes are approximately equal (17, 16, and 17 for rehearsed tasks and 17, 15, and 16 for new tasks in this

58

study), ANOVA is fairly robust even when the homogeneity of variance assumption has been violated (Field, 2005).[15]

The assumption of sphericity must also be met in repeated-measures designs and consequently also in mixed-design ANOVAs. However, for sphericity to be an issue, there must be at least three within-subject conditions (Field, 2005). In the present study, there were only two conditions for the within-subject variable (i.e., pre-test and post-test); thus the assumption of sphericity did not need to be tested.

For ANCOVA, there is an additional assumption of homogeneity of regression slopes, which was met. That is, the interaction between the group and the covariate was not significant at the $p < .05$ level for any of the dependent variables for which ANCOVAs were conducted.

For post-hoc comparison of groups, the Bonferroni's test was chosen because (a) sample sizes were practically equal; (b) this test controls the Type I error rate well; and (c) this test is powerful when the number of comparisons is small (three in the present study).

In addition to inferential statistics, qualitative analyses were also conducted. To answer research question 3 about participants' attitudes towards CASTs and role-plays in pairs, qualitative analyses and descriptive statistics were applied to the Exit Questionnaire. Although not part of the research questions, some additional qualitative analyses of the data were conducted to aid the interpretation of findings and to triangulate the data. Specifically, when rating the responses in terms of pragmatic appropriateness and when calculating the fluency, accuracy, and complexity measures, I observed certain trends; these will be reported on in chapter 5. Additionally, discourse patterns during the use of CASTs and during role-plays in

---

[15] One solution to non-normal distribution is the use of non-parametric tests. However, there is no non-parametric equivalent to mixed-design ANOVA.

dyads are provided in chapter 5 in order to show the processes that the learners were involved in during these activities.

CHAPTER 4: RESULTS

4.1 Research Question 1

What is the effect of task rehearsal on improvements in pragmatics and speaking skills on the

same (rehearsed) task?

This research question investigates the immediate effect of rehearsal on task performance.

Mean scores on the rehearsed task (i.e., the scores averaged across all rehearsed tasks) were

calculated for eight outcome measures: pragmatic appropriateness, number of syllables per

minute, number of pruned syllables per minute, number of pauses per minute, percentage of

pause time, percentage of error-free clauses, number of clauses per AS unit, and Guiraud's Index.

The means and standard deviations are provided by group in Table 6. One trend that stands out

from the descriptive statistics is that standard deviations in the control group for pragmatics and

speech rate measures were much higher than standard deviations in the experimental groups. For

example, on the pruned syllables per minute measure on the pre-test, standard deviations for the

experimental groups were 20.69 (structured task group) and 19.47 (pair-work group), while in

the control group, the standard deviation was two times higher: 42.91.That is, there was much

more variability in the control group than in the other two groups on these measures. This fact

should be kept in mind when the experimental groups and the control group are compared

throughout the study.

Mixed-design ANOVAs, with time (pre-test, post-test) as the within-subject independent

variable and group (structured task, pair-work, control) as the between-subject independent

variable, were conducted for eight dependent variables. To examine whether there were any

group differences before the treatment and whether ANCOVAs needed to be utilized (with pre-

test scores as covariates), one-way ANOVAs were conducted on the pre-test scores for each of

Table 6: Descriptive Statistics for Averaged Scores on the Rehearsed Tasks by Group

Dependent Variable	Test	Group					
		Structured Task N = 17		Pair-Work N = 16		Control N = 17	
		M	SD	M	SD	M	SD
Pragmatic Ratings	Pre	4.13	0.60	3.94	0.58	3.78	0.82
	Post	4.37	0.72	3.89	0.55	3.71	0.98
N Syllables / Min.	Pre	179.89	21.41	167.31	17.86	191.68	35.44
	Post	180.76	22.54	169.27	16.60	184.43	31.40
N Pruned Syllables / Min.	Pre	155.95	20.69	146.52	19.47	169.65	42.91
	Post	157.91	22.34	147.68	19.67	162.84	35.61
N Pauses / Min.	Pre	1.58	1.35	2.78	1.69	1.12	1.66
	Post	1.83	1.53	3.16	2.12	1.17	1.49
% Pause time	Pre	3.20	2.93	6.19	4.12	2.35	3.43
	Post	3.75	3.32	7.15	4.56	2.74	3.75
% Error-free Clauses	Pre	59.39	11.35	62.09	10.51	68.71	12.09
	Post	55.73	8.82	63.46	9.39	69.91	9.96
N Clauses / AS-unit (log)	Pre	1.26	0.11	1.35	0.13	1.21	0.16
	Post	1.32	0.10	1.33	0.17	1.17	0.13
Guiraud's Index	Pre	5.24	0.45	5.11	0.58	4.74	0.38
	Post	5.57	0.33	5.64	0.49	5.09	0.37

the eight dependent variables. The results of one-way ANOVAs are provided in Table 7. These results indicate that at the alpha level of .006, established after the Bonferroni correction, the groups did not differ on the pre-test scores for any of the dependent variables. Thus, mixed-design ANOVAs will not be followed up with one-way ANCOVAs. In the sections that follow, the results of mixed-design ANOVAs are provided for each dependent variable.

4.1.1 Pragmatic Appropriateness on Rehearsed Tasks

As shown in Table 8, there were no significant main effects of time, group, or Time x Group interaction for pragmatic appropriateness on the rehearsed tasks. That is, there were no significant improvements between the pre-test and the post-test for all groups combined. There were also no statistically significant differences between the three groups on gains from the pre-test to the post-test. Likewise, on the pre-test and post-test combined, the three groups did not differ significantly on pragmatic appropriateness ratings. That is, the groups performed similarly

on the pre-test and on the post-test. (The readers are reminded that the alpha level of .006 is used

to determine statistical significance because due to a large number of tests performed on the

same data set, the Bonferroni correction was made).

Table 7: Results of One-Way ANOVAs on the Rehearsed Task Pre-Tests

Test	Source	df	F	p
Pragmatic Appropriateness	Between Groups	2	1.117	.336
	Within Groups (Error)	47	(.461)	
N Syllables / Minute	Between Groups	2	3.569	.036
	Within Groups (Error)	47	(685.361)	
N Pruned Syllables / Minute	Between Groups	2	2.505	.093
	Within Groups (Error)	47	(893.586)	
N Pauses / Minute	Between Groups	2	4.804	.013
	Within Groups (Error)	47	(2.475)	
% Pause time	Between Groups	2	5.381	.008
	Within Groups (Error)	47	(12.338)	
% Error-free Clauses	Between Groups	2	3.023	.058
	Within Groups (Error)	47	(128.834)	
N Clauses / AS-unit (log)	Between Groups	2	5.163	.009
	Within Groups (Error)	47	(.002)	
Guiraud's Index	Between Groups	2	5.094	.010
	Within Groups (Error)	47	(.224)	

Note. α = .006. Values enclosed in parentheses represent mean square errors.

Table 8: Pragmatic Appropriateness on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	0.100	.002	.754
Time x Group	2	0.674	.028	.514
Error	47	(.396)		
Between subjects				
Group	2	3.380	.126	.043
Error	47	(.668)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.

The results on pragmatic appropriateness are graphically represented in Figure 3. Based

on descriptive statistics, the trends follow the hypothesis: the structured task group improved

(from $M = 4.13$ on the pre-test to $M = 4.37$ on the post-test), while the control group did not.

However, contrary to the prediction, the pair-work group did not improve even though it

received instruction on pragmatics. The reasons for the lack of statistically significant

improvement on pragmatics are examined in chapter 5.[16]
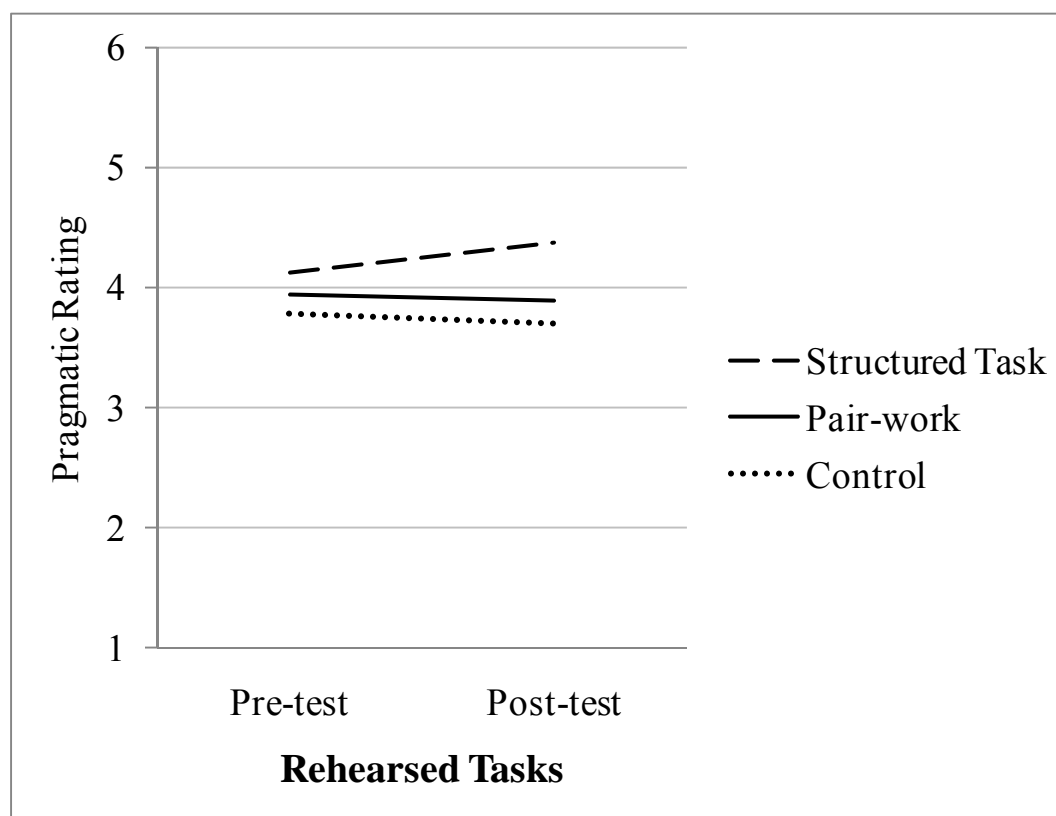
Figure 3. Pragmatic Appropriateness Ratings on the Rehearsed Tasks

4.1.2 Fluency on Rehearsed Tasks

Fluency was examined via four measures: number of syllables per minute, number of

pruned syllables per minute, number of pauses per minute, and percentage of pause time.

---

[16] Recall that there were three responses with a pragmatic rating of 1 (extremely inappropriate) on rehearsed tasks, and qualitatively these responses looked like outliers. A follow-up analysis was conducted, in which participants who received these ratings were excluded. However, the results did not change: the relationship between groups remained the same, and main effects of time, group, and Time x Group interaction were not significant.

For the syllables per minute measure, the main effects of time and group and the Time x Group interaction were non-significant (see Table 9). That is, there were no significant improvements between the pre-test and the post-test for all groups combined; the groups also did not differ on gains between the pre-test and the post-test, as well as on the combined pre-test and post-test performance. As shown in Figure 4, the trend was that the control group had the highest speech rate. This trend persists for other measures of fluency on rehearsed tasks, and will be examined in the discussion section.

Table 9: Syllables per Minute on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	.479	.010	.492
Time x Group	2	1.876	.074	.165
Error	47	(112.996)		
Between subjects				
Group	2	2.786	.106	.072
Error	47	(1168.933)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.
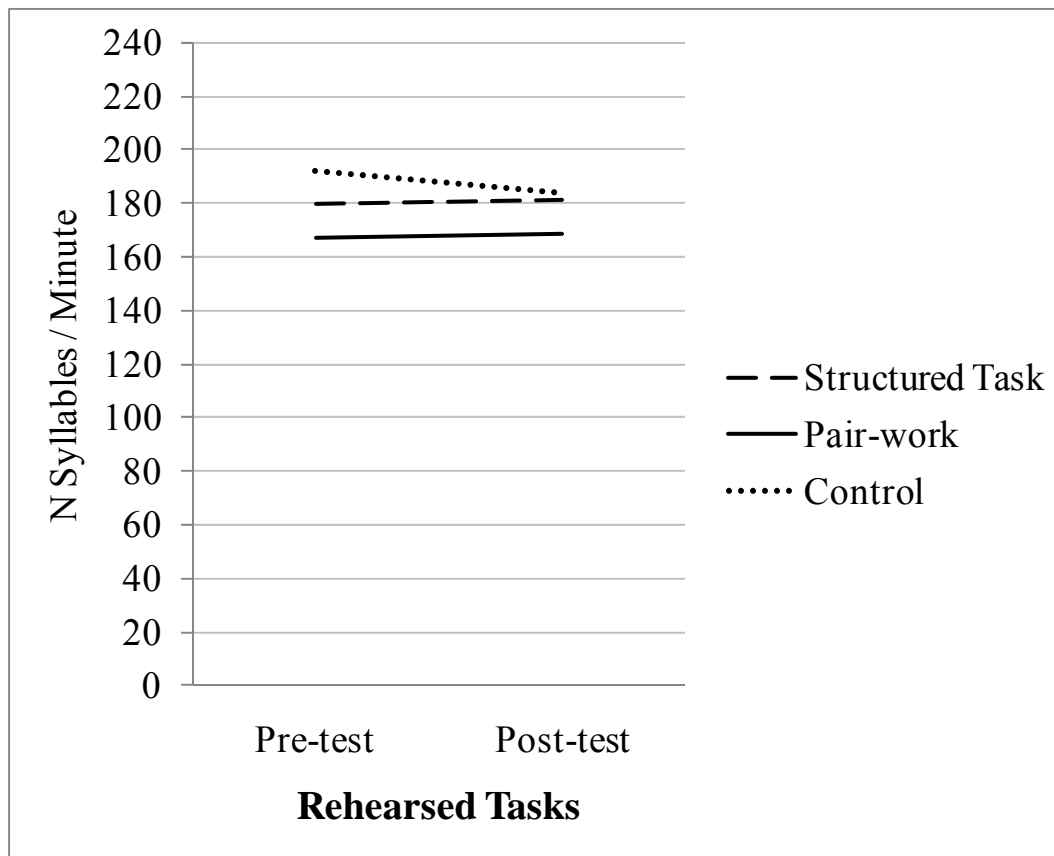
Figure 4. Number of Syllables per Minute on the Rehearsed Tasks

In addition to syllables per minute, pruned syllables per minute was used as a measure of meaningful production in that disfluencies, such as false starts and repetitions, were not included in the syllable count. However, this measure produced the same results as number of syllables per minute both in terms of a mixed-design ANOVA and trends. That is, the data yielded no significant main effects of time or group and a non-significant Time x Group interaction (see Table 10), and the trend was that the control group had the highest speech rate (see Figure 5).

Table 10: Pruned Syllables per Minute on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	.317	.007	.576
Time x Group	2	1.669	.066	.199
Error	47	(118.652)		
Between subjects				
Group	2	2.014	.079	.145
Error	47	(1500.033)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.
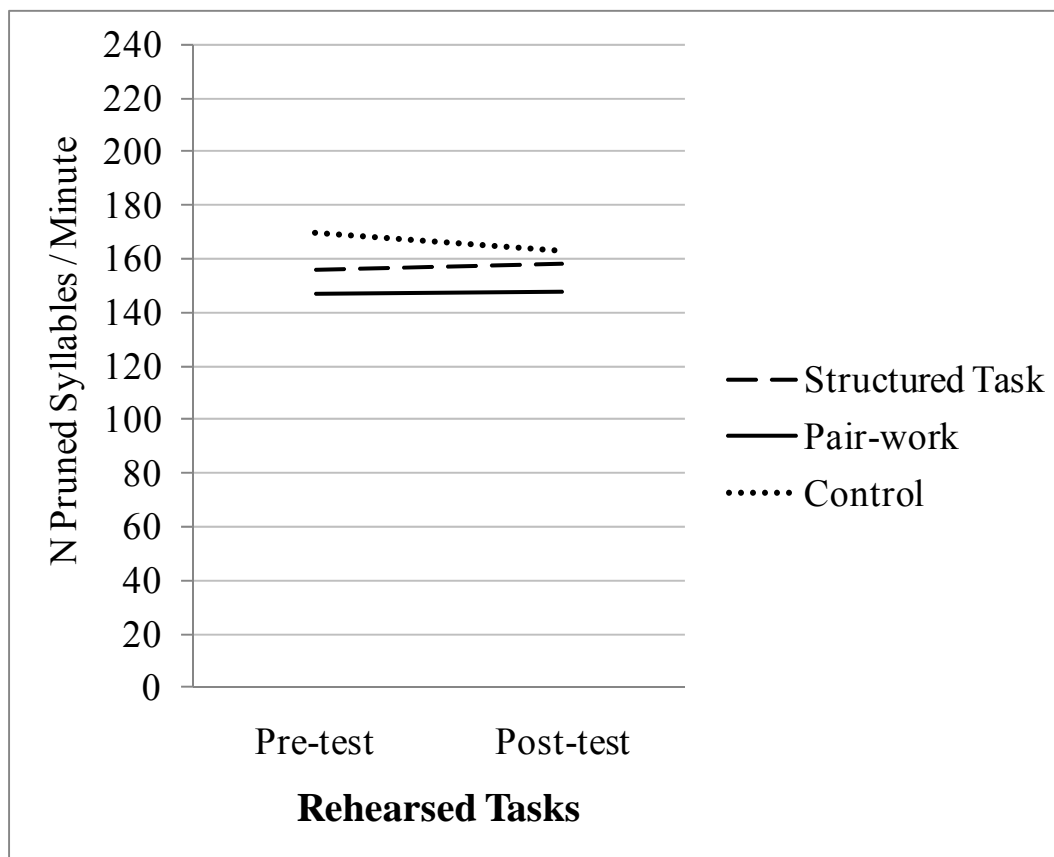
Figure 5. Number of Pruned Syllables per Minute on the Rehearsed Tasks

Pausing phenomena may also contribute to fluency, and thus number of pauses per minute and percentage of pausing time were examined. As shown in Table 11, there was no significant main effect of time and no significant Time x Group interaction. That is, there was no significant change between the pre-test and the post-test for all groups combined, and the groups

also did not differ in gains. However, there was a significant main effect of group ($p = .005$,

$\eta_p^2 = .204$). The post-hoc Bonferroni test showed that the control group produced significantly

fewer pauses per minute than the pair-work group ($p = .004$). The results are graphically

represented in Figure 6.

Table 11: Number of Pauses per Minute on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	1.752	.036	.192
Time x Group	2	.323	.014	.725
Error	47	(.746)		
Between subjects				
Group	2	6.036	.204	.005
Error	47	(4.720)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.
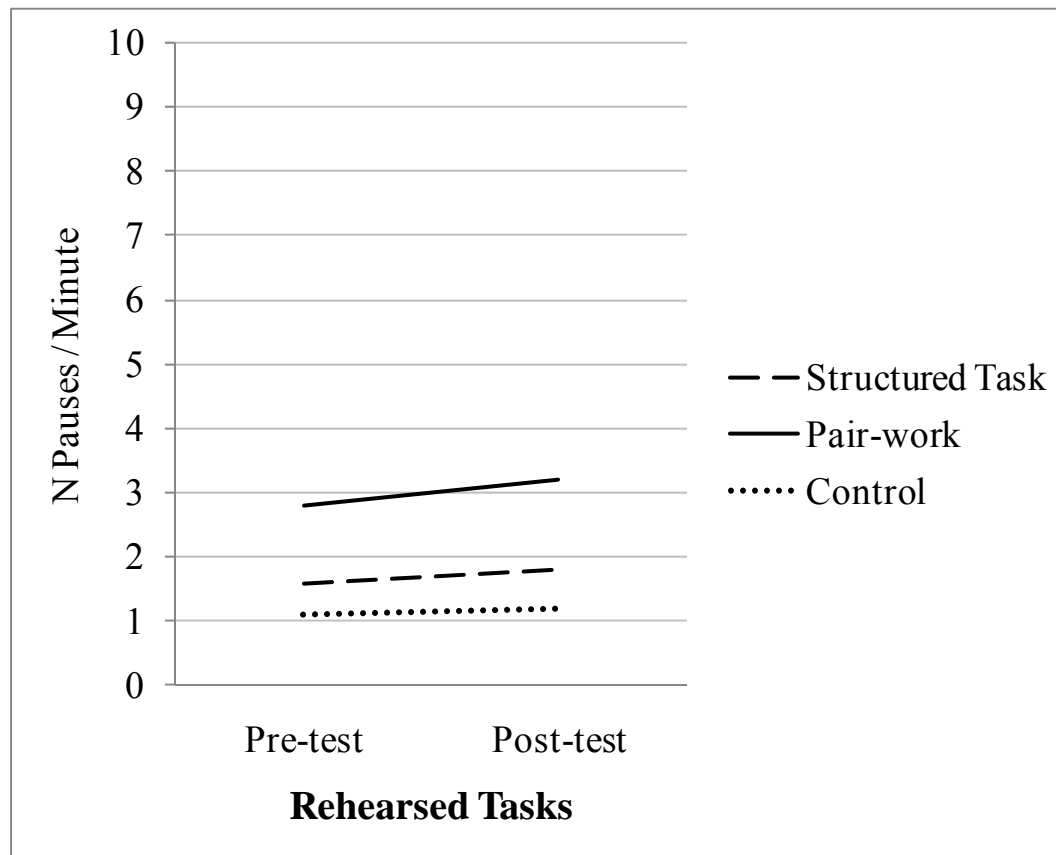
Figure 6. Number of Pauses per Minute on the Rehearsed Tasks

The percentage of pausing time measure produced results similar to the number of pauses per minute. As shown in Table 12, there was no significant main effect of time and no significant Time x Group interaction. That is, there was no significant change between the pre-test and the post-test for all groups combined, and the groups also did not differ in gains. However, there was a significant main effect of group ($p = .003$, $\eta_p^2 = .218$). The post-hoc Bonferroni test showed that the pair-work group spent a larger percentage of time pausing than the control group ($p = .003$) and the structured task group ($p = .030$). The results are graphically represented in Figure 7.

Table 12: Percentage of Pausing Time on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	2.422	.049	.126
Time x Group	2	.176	.007	.839
Error	47	(4.092)		
Between subjects				
Group	2	6.532	.218	.003
Error	47	(23.420)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.

A common result for all fluency measures is that there were no improvements on fluency, and the control group was the most fluent. However, the reader is also reminded that the results on fluency measures need to be interpreted with caution because the assumption of homogeneity of variances was not met for speech rate measures and the assumption of normal distribution was not met for pause-related measures.

4.1.3 Accuracy on Rehearsed Tasks

Accuracy was measured as a percentage of error-free clauses. As shown in Table 13, there was no significant main effect of time and no significant Time x Group interaction. That is, there was no significant change between the pre-test and the post-test for all groups combined,

and the groups also did not differ in gains. However, there was a significant main effect of group ($p = .002$, $\eta_p^2 = .233$). The post-hoc Bonferroni test showed that the control group had a significantly higher percentage of error-free clauses than the structured task group ($p = .001$).
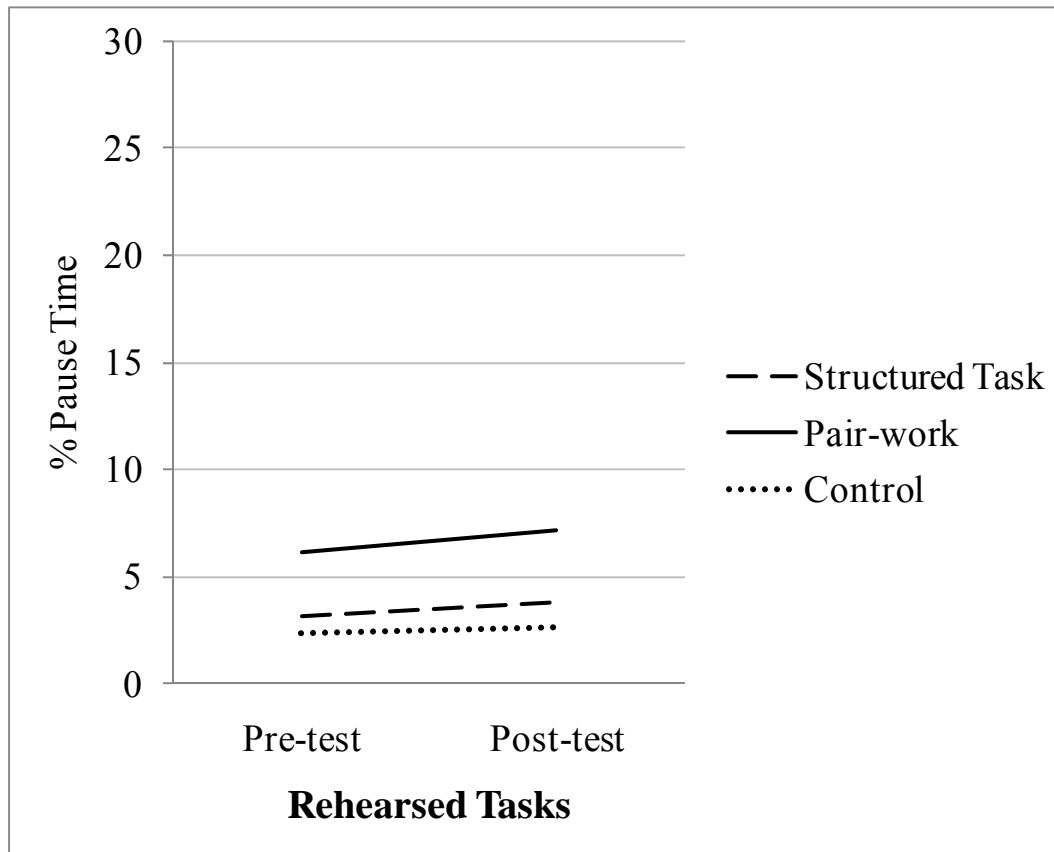
Figure 7. Percentage of Pausing Time on the Rehearsed Tasks

Table 13: Percentage of Error-free Clauses on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	.064	.001	.802
Time x Group	2	1.311	.053	.279
Error	47	(52.445)		
Between subjects				
Group	2	7.149	.233	.002
Error	47	(164.819)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.

The trend is that the control group produced more accurate responses than both experimental groups (see Figure 8).
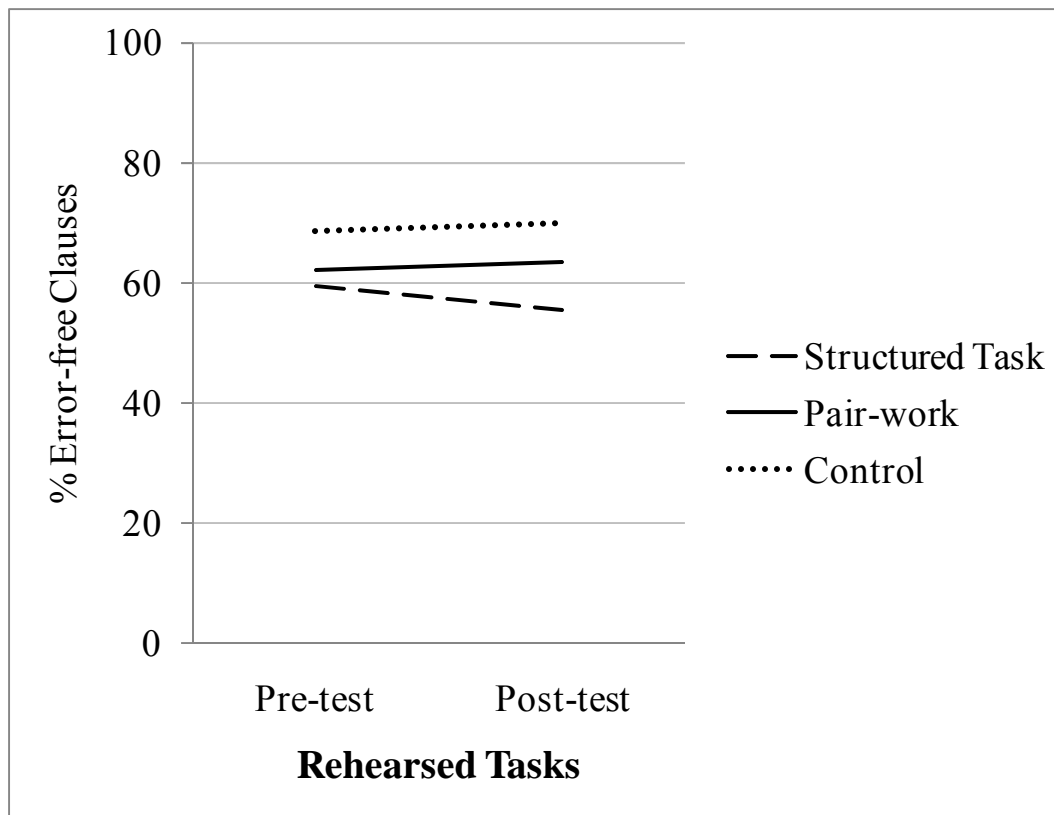
Figure 8. Percentage of Error-free Clauses on the Rehearsed Tasks

4.1.4 Complexity on Rehearsed Tasks

Two measures of complexity were examined: syntactic complexity (number of clauses per AS-unit) and vocabulary complexity / diversity (Guiraud's Index). For syntactic complexity, there was no significant main effect of time and no significant Time x Group interaction (see Table 14). That is, there was no significant change between the pre-test and the post-test for all groups combined, and the groups also did not differ in gains. However, there was a significant main effect of group ($p < .001$, $\eta_p^2 = .283$). The post-hoc Bonferroni test showed that the control group produced fewer clauses per AS-unit than the structured task group ($p = .013$) and the pair-work group ($p < .001$). That is, control group's responses were the least syntactically complex, as

71

can be seen in Figure 9. This finding is different from previously mentioned results where the

control group produced more accurate responses and was more fluent than the experimental

groups, but this is also understandable given the inverse relationship between fluency/accuracy

and complexity.

Table 14: Number of Clauses per AS-unit (log) on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	.007	< .001	.934
Time x Group	2	1.609	.064	.211
Error	47	(.002)		
Between subjects				
Group	2	9.298	.283	< .001
Error	47	(.003)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
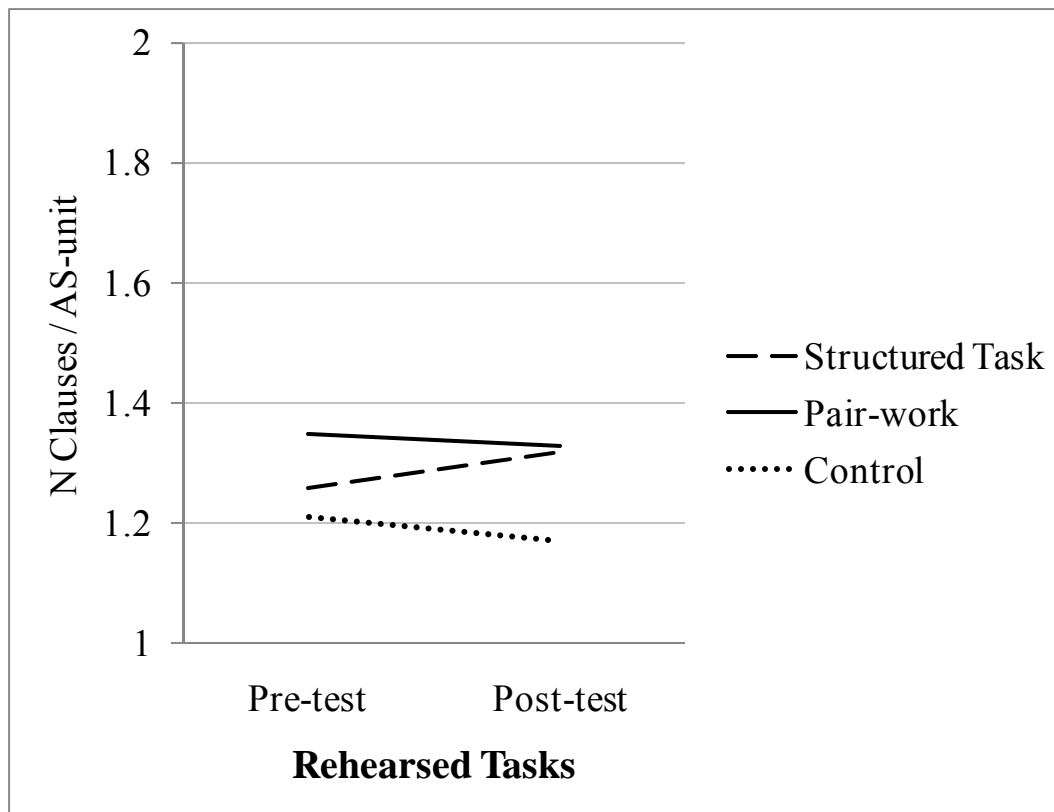
Figure 9. Number of Clauses per AS-unit on the Rehearsed Tasks

For vocabulary diversity, as shown in Table 15, there was a significant main effect of time ($p < .001$), and according to Cohen's (1988) guidelines for effect sizes measured in eta squared (.01, .059, and .138 for small, medium, and large effects respectively), the effect size was large ($\eta_p^2 = .552$). That is, for vocabulary diversity there was a significant improvement between the pre-test and the post-test for all groups combined. The groups did not differ on gains as there was no significant Time x Group interaction, but the main effect of group was significant ($p = .001$, $\eta_p^2 = .258$). The post-hoc Bonferroni test showed that the control group's responses were lower on vocabulary diversity than those of the structured task group ($p = .002$) and of the pair-work group ($p = .005$) (also see Figure 10). Thus, control group's responses were lowest on both measures of complexity.

Table 15: Guiraud's Index on the Rehearsed Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	57.952	.552	< .001
Time x Group	2	1.257	.051	.294
Error	47	(.071)		
Between subjects				
Group	2	8.151	.258	.001
Error	47	(.312)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.

4.1.5 Summary of Mixed-design ANOVA Results on Rehearsed Tasks

A summary of the results of eight mixed-design ANOVAs on the rehearsed tasks is provided in Table 16. First, the main effect of time was significant only for Guiraud's Index. That is, the responses of the three groups combined improved only on vocabulary diversity, although it was predicted that there would be improvements on other measures, in particular on pragmatics. None of the Time x Group interactions was significant, indicating that the groups did

not differ on gains. However, descriptive statistics show that, as predicted, the structured task group improved on pragmatics while the control group did not.
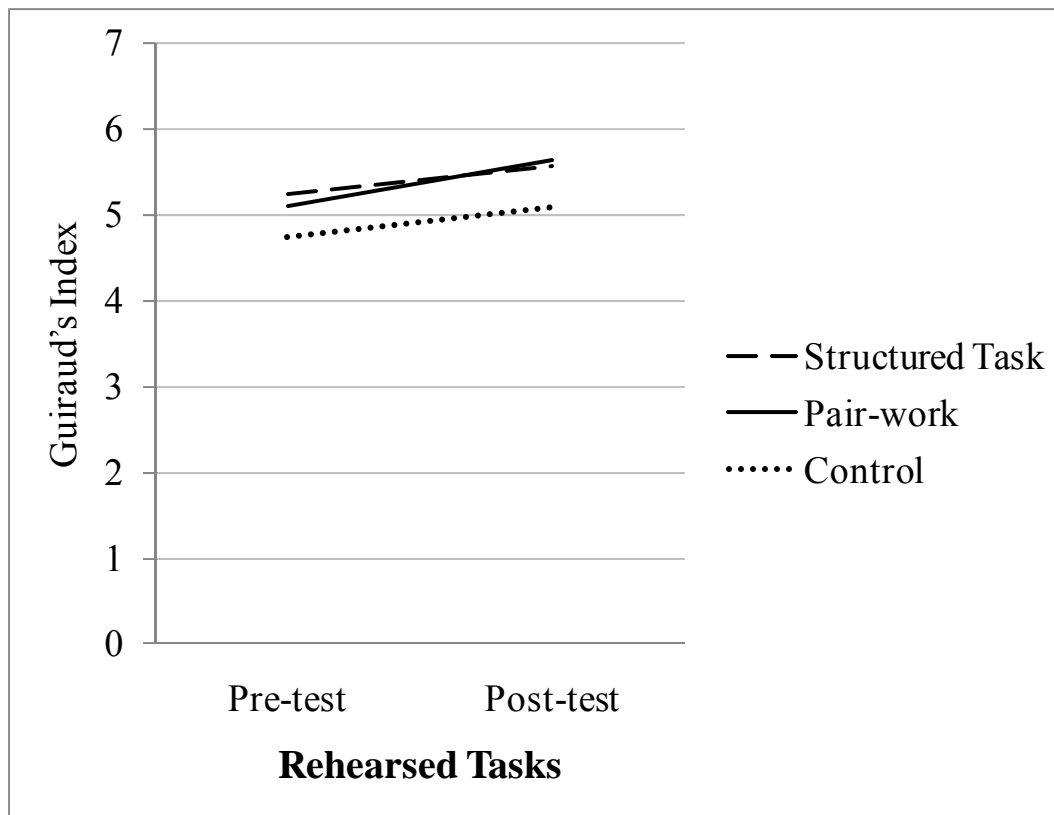
Figure 10. Guiraud's Index on the Rehearsed Tasks

Table 16: Summary of Effect Sizes of Mixed-design ANOVA Results on the Rehearsed Tasks

Dependent Variable	Main Effect of Time	Time x Group Interaction	Main Effect of Group	Group Differences
Pragmatic Appropriateness	.002	.028	.126	
N Syllables / Minute	.010	.074	.106	
N Pruned Syllables / Minute	.007	.066	.079	
N Pauses / Minute	.036	.014	.204*	C < P
% Pause time	.049	.007	.218*	C < P, S < P
% Error-free Clauses	.001	.053	.233*	C > S
N Clauses / AS-unit (log)	< .001	.064	.283*	C < S, P
Guiraud's Index	.552*	.051	.258*	C < S, P

Note. The numbers represent effect sizes. S = structured task group. P = pair-work group. C = control group.
* p < .006

74

A major finding was that the main effect of group was significant for five out of eight dependent variables. That is, while there were no statistically significant differences between the groups on the pre-test at the alpha level of .006, these differences surfaced when performance on both pre-tests and post-tests was taken into account. In only one case, the two experimental groups differed: the pair-work group had a higher percentage of pausing time than the structured task group. In all other cases, the differences were between the control group and one or both of the experimental groups. (This goes along with the observation pointed out earlier that, judging by standard deviations, there was more variability within the control group than within the experimental groups). The control group's responses were lower than those of the experimental groups on syntactic complexity and vocabulary diversity. Interestingly, the control group produced language that was more accurate than that of the structured task group. The control group made fewer pauses per minute and paused for a shorter percentage of time than the pair-work group. The effect sizes for these group differences were large, with partial eta squared ranging from .204 to .283. While there were no statistically significant differences between the three groups on the two speech rate measures, the control group had the highest speech rates. Thus, the four measures of fluency seem to indicate that the control group was the most fluent. The fact that the groups differed in the same way on the pre-test and on the post-test suggests that group differences may be due to factors other than treatment, and possible reasons are examined closely in the discussion section.

4.2 Research Question 2

To what extent can task rehearsal influence performance on a new task?

This research question investigates whether the effect of task rehearsal is generalizable to new tasks (i.e., to language development in general). The mean scores on the new task (i.e., the

scores averaged across the three new tasks) were calculated for the same eight outcome measures as for research question 1: pragmatic appropriateness, number of syllables per minute, number of pruned syllables per minute, number of pauses per minute, percentage of pause time, percentage of error-free clauses, number of clauses per AS unit, and Guiraud's Index. The means and standard deviations are provided in Table 17.[17] For new tasks, standard deviations in the control group are also much higher than those in the two experimental groups.

Next, the same statistical tests were conducted as for research question 1. Mixed-design ANOVAs, with time (pre-test, post-test) as the within-subject independent variable and group (structured task, pair-work, control) as the between-subject independent variable, were conducted for eight dependent variables. To examine whether there were any group differences before the treatment and whether ANCOVAs needed to be utilized (with pre-test scores as a covariate), one-way ANOVAs were conducted on the pre-test scores for each of the eight dependent variables. The results of one-way ANOVAs are provided in Table 18. These results indicate that at the alpha level of .006, established after the Bonferroni correction, the groups differed only in terms of Guiraud's Index: the control group had lower vocabulary diversity than the experimental groups. Thus, only the mixed-design ANOVA for vocabulary

---

[17] For various measures, there appear to be differences between the pre-test scores on the rehearsed task and the pre-test scores on the new task, although statistical tests were not performed to investigate whether these differences are statistically significant. These differences are probably due to the different formats of the rehearsed task pre-test (role-plays in pairs) and the new task pre-test (multi-turn oral DCTs conducted via CASTs). The interactive nature of role-plays in pairs can affect the calculation of different measures. For example, recall that the AS-unit, introduced by Foster et al. (2000) and used in this study, allows for a slight modification of oral highly-interactive data, such as deletion of one-word minor utterances and one-word elliptical utterances, echo responses which are verbatim, and one- or two-word greetings and closures. On the other hand, multi-turn oral DCTs are less interactive, and speech samples obtained from them contain longer stretches of discourse.

diversity will be followed up with a one-way ANCOVA. In the sections that follow, the results

of mixed-design ANOVAs are provided first for each dependent variable.

Table 17: Descriptive Statistics for Averaged Scores on New Tasks by Group

Dependent Variable	Test	Group					
		Structured Task $N = 17$		Pair-Work $N = 15$		Control $N = 16$	
		M	SD	M	SD	M	SD
Pragmatic Ratings	Pre	4.13	0.77	3.78	0.70	3.53	0.90
	Post	4.36	0.64	4.49	0.72	3.92	0.72
N Syllables / Minute	Pre	165.54	18.72	159.06	16.03	174.95	36.29
	Post	165.80	19.62	164.33	16.25	179.16	28.08
N Pruned Syllables / Min.	Pre	142.03	19.91	140.42	18.47	157.66	40.05
	Post	141.72	22.54	144.40	19.76	162.22	31.44
N Pauses / Min.	Pre	2.12	1.46	3.28	2.00	1.71	2.17
	Post	2.44	1.66	3.67	2.15	2.69	2.80
% Pause time	Pre	4.61	3.22	7.51	5.12	4.72	7.18
	Post	4.98	3.56	7.91	4.26	5.75	6.70
% Error-free Clauses	Pre	59.32	9.12	65.20	11.15	69.56	16.34
	Post	60.70	9.94	69.02	12.64	68.09	15.41
N Clauses / AS-unit (log)	Pre	1.24	0.09	1.24	0.16	1.17	0.12
	Post	1.30	0.09	1.27	0.13	1.21	0.12
Guiraud's Index	Pre	5.21	0.45	5.25	0.53	4.54	0.55
	Post	5.46	0.40	5.31	0.62	4.79	0.65

Table 18: Results of One-Way ANOVAs on the New Task Pre-Tests

Test	Source	df	F	p
Pragmatic Appropriateness	Between Groups	2	2.346	.107
	Within Groups (Error)	45	(.632)	
N Syllables / Minute	Between Groups	2	1.543	.225
	Within Groups (Error)	45	(643.492)	
N Pruned Syllables / Minute	Between Groups	2	1.844	.170
	Within Groups (Error)	45	(781.753)	
N Pauses / Minute	Between Groups	2	2.862	.068
	Within Groups (Error)	45	(3.572)	
% Pause time	Between Groups	2	1.445	.246
	Within Groups (Error)	45	(28.999)	
% Error-free Clauses	Between Groups	2	2.775	.073
	Within Groups (Error)	45	(157.249)	
N Clauses / AS-unit (log)	Between Groups	2	2.105	.134
	Within Groups (Error)	45	(.002)	
Guiraud's Index	Between Groups	2	9.949	< .001
	Within Groups (Error)	45	(.259)	

Note. α = .006. Values enclosed in parentheses represent mean square errors.

4.2.1 Pragmatic Appropriateness on New Tasks

While on the rehearsed tasks the main effect of time was not significant for pragmatics (i.e., there were no significant improvements on pragmatics between the pre-test and the post-test for all groups combined), the results were different for the new tasks. As shown in Table 19, there was a significant main effect of time ($p < .001$), and the effect size was large (η_p^2 = .382).

The Time x Group interaction was not significant; that is, the groups did not differ on gains between the pre-test and the post-test. These results contradict the hypothesis that the control group would not improve on pragmatics since it did not receive instruction in this area. It was also predicted that the structured task group would improve more than the pair-work group, which was not borne out. There was no significant main effect of group, which is the same result as the one obtained for rehearsed tasks. The results are graphically represented in Figure 11.

Table 19: Pragmatic Appropriateness on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	27.865	.382	< .001
Time x Group	2	2.760	.109	.074
Error	45	(.172)		
Between subjects				
Group	2	2.588	.103	.086
Error	45	(.937)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
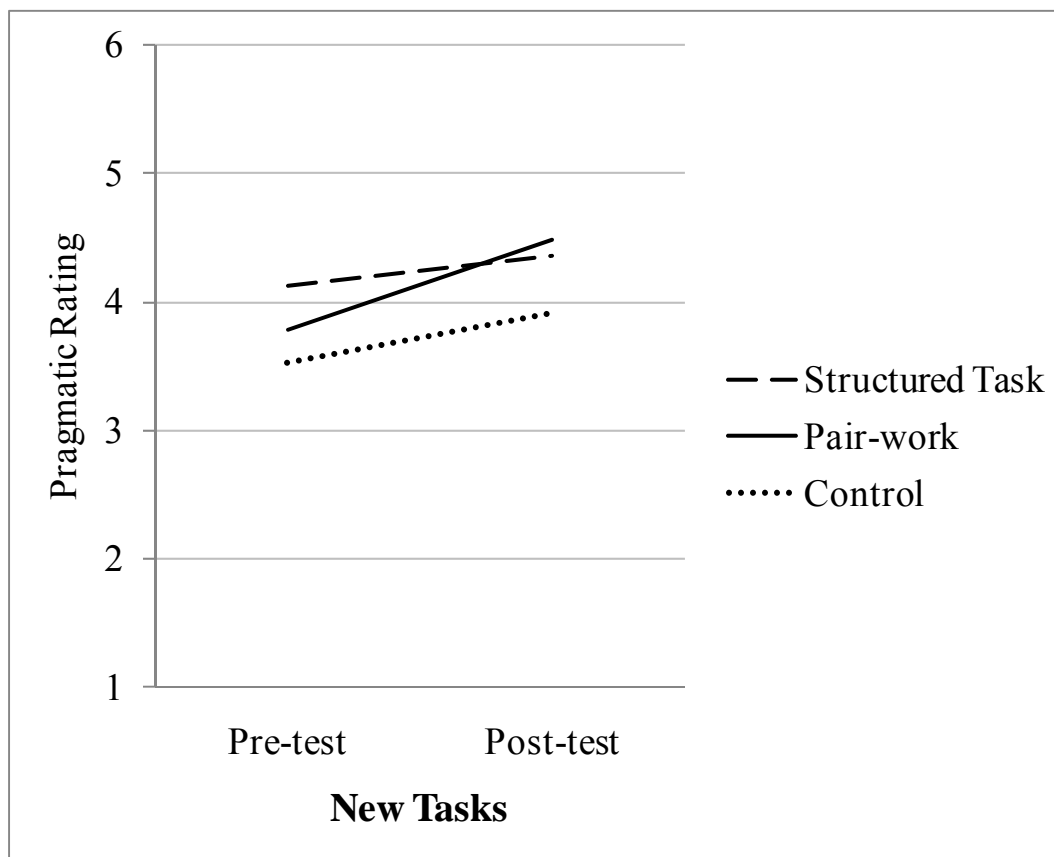
Figure 11. Pragmatic Appropriateness Ratings on the New Tasks

4.2.2 Fluency on New Tasks

Just like for research question 1, fluency was examined via four measures: number of syllables per minute, number of pruned syllables per minute, number of pauses per minute, and percentage of pause time.

For the syllables per minute measure, the results were the same as for rehearsed tasks. There were no significant main effects of time and group and no significant Time x Group interaction (see Table 20). That is, there were no significant improvements between the pre-test and the post-test for all groups combined; the groups did not differ on gains between the pre-test and the post-test, as well as on combined pre-test and post-test performance. As shown in Figure 12, the trend was that the control group had the highest speech rate. The same trend surfaced in research question 1.

Table 20: Syllables per Minute on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	1.902	.041	.175
Time x Group	2	.429	.019	.654
Error	45	(132.690)		
Between subjects				
Group	2	2.011	.082	.146
Error	45	(992.708)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.

Pruned syllables per minute, a measure of meaningful production, showed the same results as the syllables per minute measure both in terms of mixed-design ANOVA and trends. That is, there were no significant main effects of time and group and no significant Time x Group interaction (see Table 21), and the trend was that the control group spoke faster than the two experimental groups (see Figure 13).

80
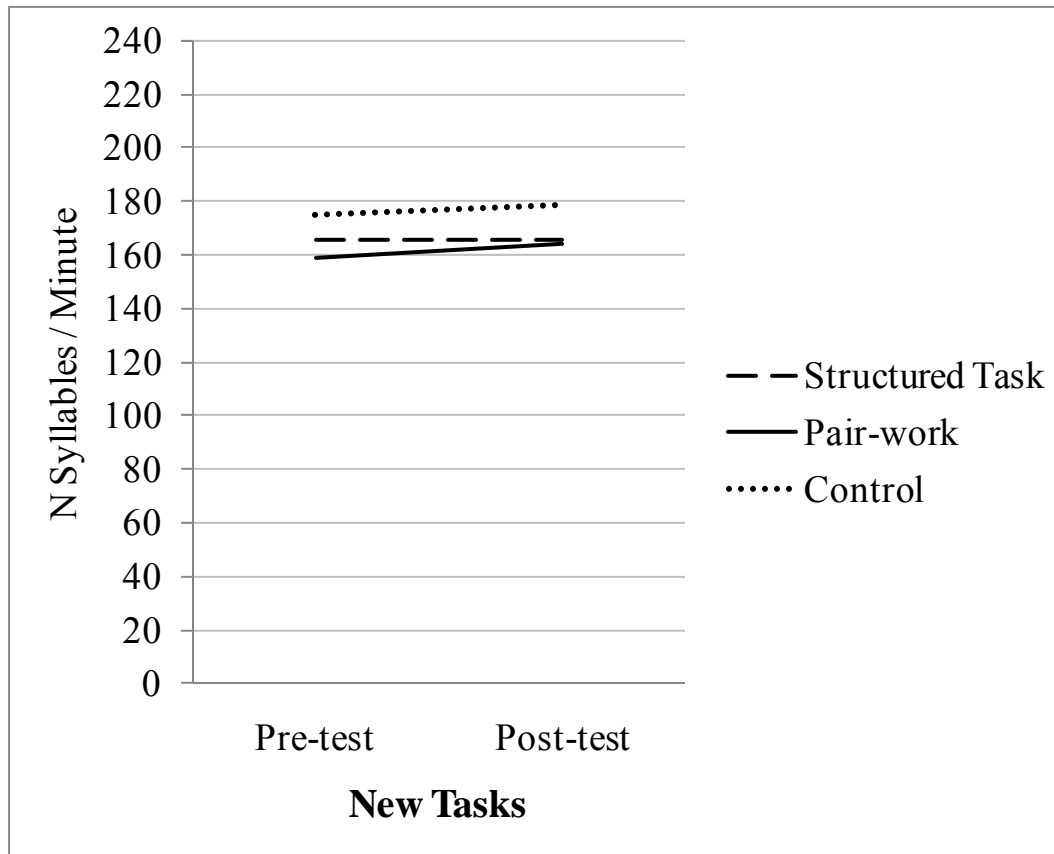
Figure 12. Number of Syllables per Minute on the New Tasks

Table 21: Pruned Syllables per Minute on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	1.095	.024	.301
Time x Group	2	.356	.016	.702
Error	45	(164.387)		
Between subjects				
Group	2	2.711	.108	.077
Error	45	(1249.022)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
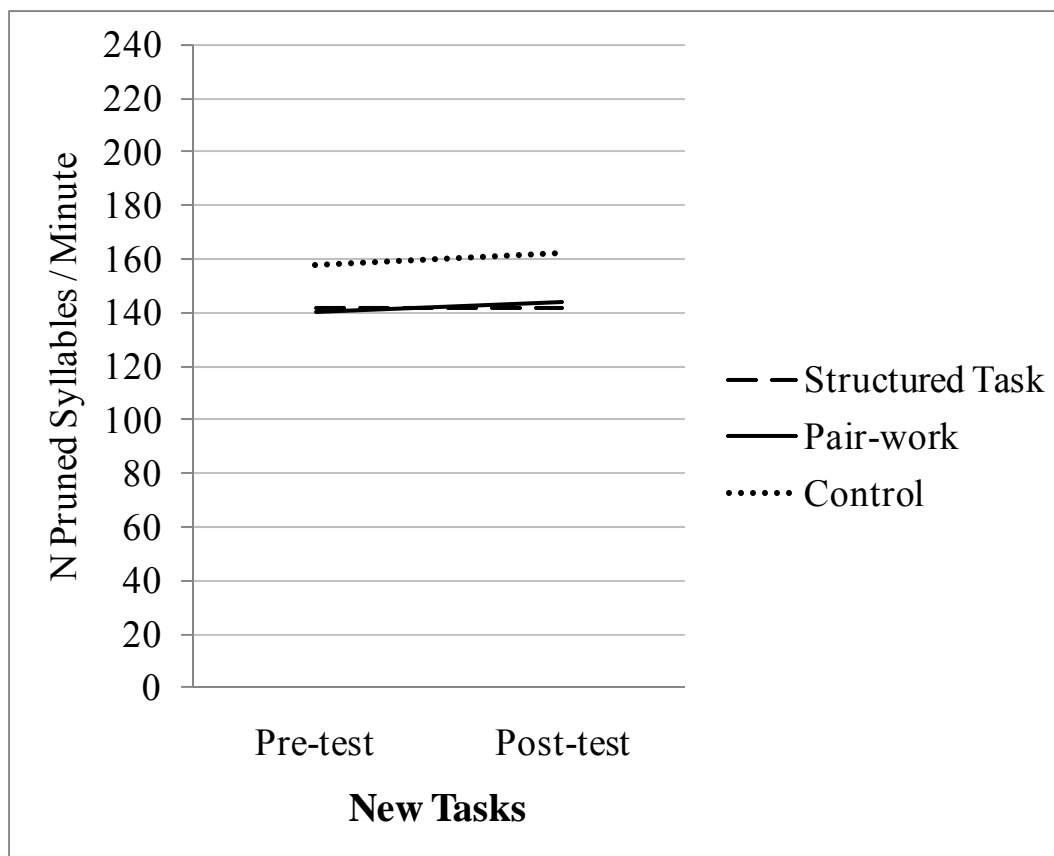
Figure 13. Number of Pruned Syllables per Minute on the New Tasks

Pausing phenomena were also examined. As shown in Table 22, at the alpha level of .006, there was no significant main effect of time and no significant Time x Group interaction (i.e., there were no changes for the number of pauses per minute between the pre-test and the post-test for all groups combined, and the groups also did not differ on gains). Additionally, as opposed to rehearsed task results, there was no significant main effect of group (i.e., the groups did not differ significantly on their combined pre-test and post-test performance). The results are graphically represented in Figure 14.

Table 22: Number of Pauses per Minute on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	5.205	.104	.027
Time x Group	2	.723	.031	.491
Error	45	(1.444)		
Between subjects				
Group	2	2.209	.089	.122
Error	45	(7.155)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
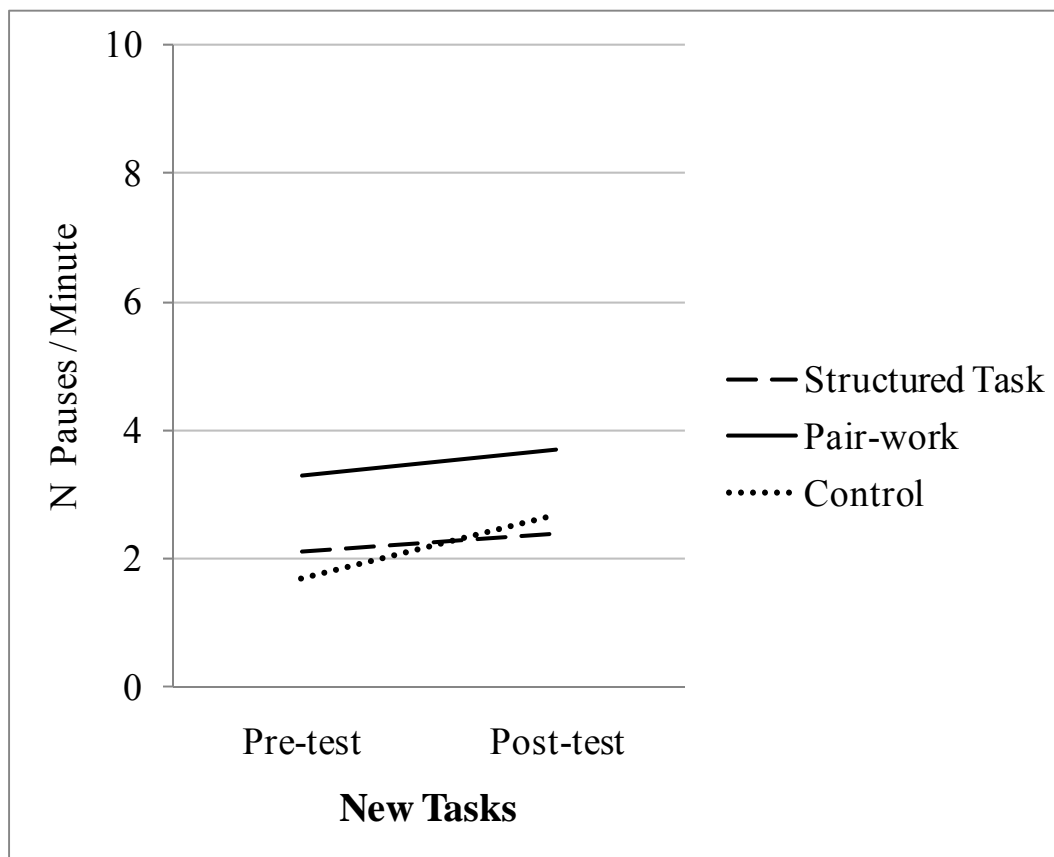
Figure 14. Number of Pauses per Minute on the New Tasks

The percentage of pausing time measure produced results similar to the number of pauses per minute. As shown in Table 23, there were no significant main effects of time and group and no significant Time x Group interaction. The results are graphically represented in Figure 15.

Table 23: Percentage of Pausing Time on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	1.114	.024	.297
Time x Group	2	.143	.006	.867
Error	45	(7.765)		
Between subjects				
Group	2	1.661	.069	.201
Error	45	(46.332)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
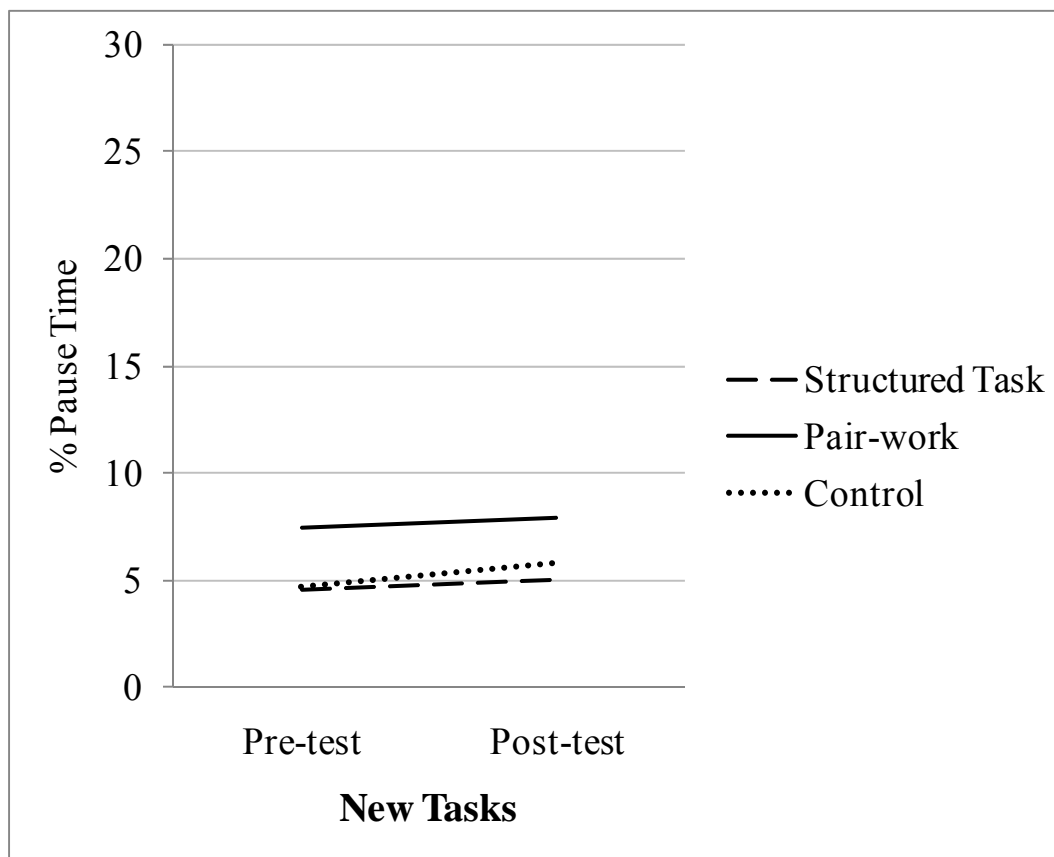
Figure 15. Percentage of Pausing Time on the New Tasks

It should be kept in mind that the results on fluency are tentative because fluency measures violated either the assumption of normal distribution or of homogeneity of variances.

4.2.3 Accuracy on New Tasks

Accuracy was measured as a percentage of error-free clauses. As shown in Table 24, there were no significant main effects of time and group and no significant Time x Group interaction. However, descriptively the control group had a higher percentage of error-free clauses than the structured task group, and the same trend was observed on rehearsed tasks. The results are graphically represented in Figure 16.

Table 24: Percentage of Error-free Clauses on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	.582	.013	.450
Time x Group	2	.857	.037	.431
Error	45	(63.488)		
Between subjects				
Group	2	2.806	.111	.071
Error	45	(257.741)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.
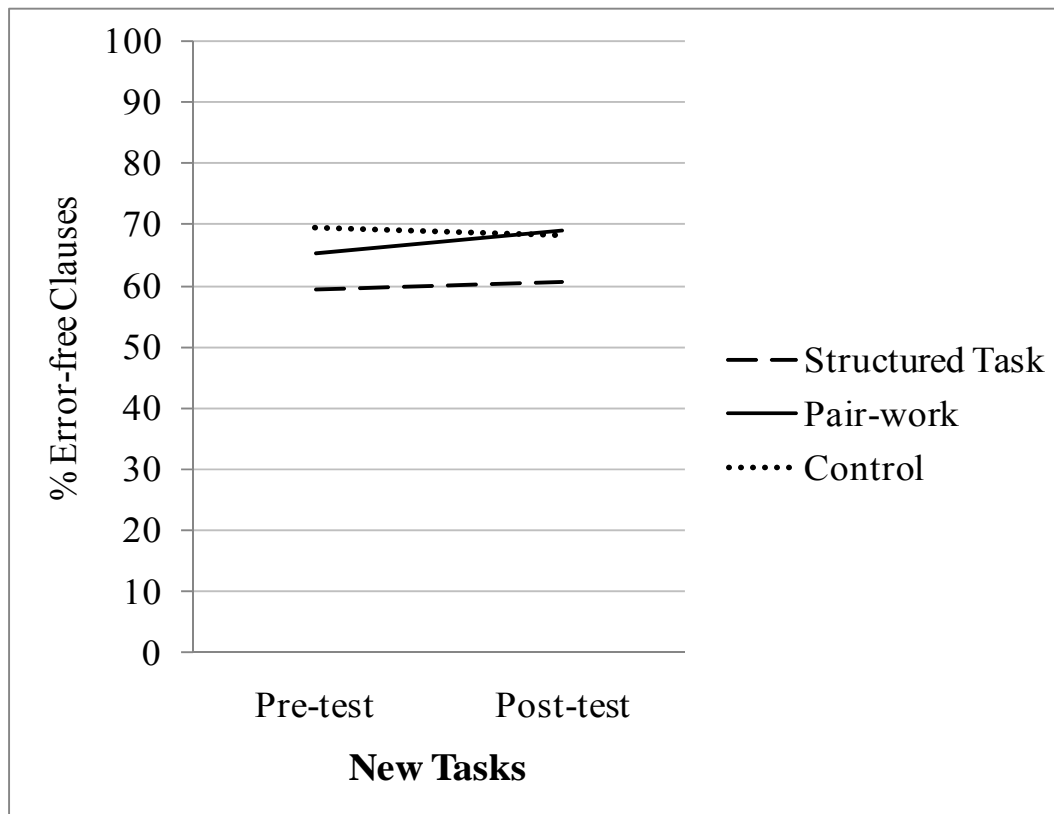
Figure 16. Percentage of Error-free Clauses on the New Tasks

4.2.4 Complexity on New Task

Two measures of complexity were examined: syntactic complexity (number of clauses per AS-unit) and vocabulary complexity / diversity (Guiraud's Index). For syntactic complexity, there were no significant main effects of time and group and no significant Time x Group interaction at the alpha level of .006 (see Table 25). That is, there was no improvement on syntactic complexity between the pre-test and the post-test, and the groups did not differ on gains and on combined pre-test and post-test performance. The results are graphically represented in Figure 17.

Table 25: Number of Clauses per AS-unit (log) on the New Tasks

Source	*df*	*F*	η_p^2	*p*
	Within subjects			
Time	1	5.127	.102	.028
Time x Group	2	.113	.005	.894
Error	45	(.001)		
	Between subjects			
Group	2	3.313	.128	.045
Error	45	(.003)		

Note. $\alpha = .006$. Values enclosed in parentheses represent mean square errors.
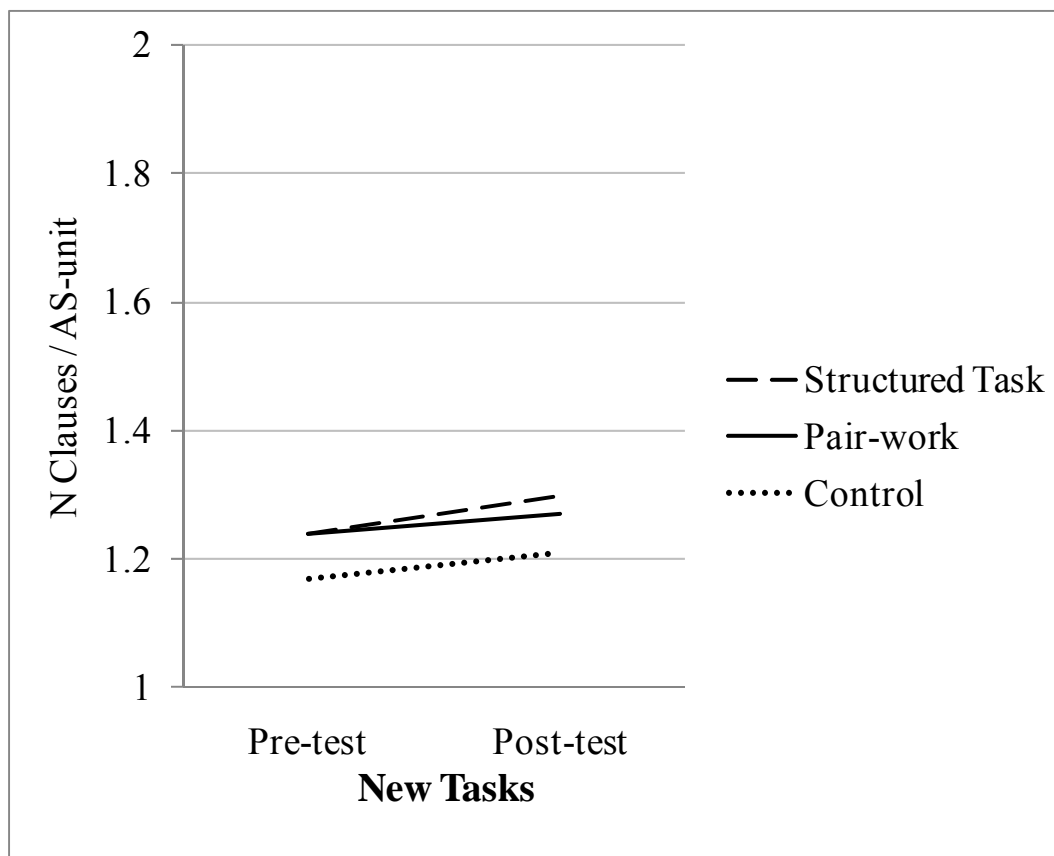
Figure 17. Number of Clauses per AS-unit on the New Tasks

For vocabulary diversity, as shown in Table 26, there was a significant main effect of time ($p = .004$), and according to Cohen's (1988) guidelines, the effect size was large ($\eta_p^2 = .173$). That is, there was a statistically significant improvement between the pre-test and the post-test for all groups combined. There was also a significant main effect of group ($p < .001$, $\eta_p^2 = .292$). The post-hoc Bonferroni test showed that the control group's responses were lower on vocabulary diversity than those of the structured task group ($p = .001$) and of the pair-work group ($p = .003$), which is the same result as the one on research question 1. However, the Time x Group interaction was not significant (i.e., the groups did not differ on gains).The results are graphed in Figure 18.

Table 26: Guiraud's Index on the New Tasks

Source	df	F	η_p^2	p
Within subjects				
Time	1	9.435	.173	.004
Time x Group	2	1.033	.044	.364
Error	45	(.088)		
Between subjects				
Group	2	9.299	.292	< .001
Error	45	(.486)		

Note. α = .006. Values enclosed in parentheses represent mean square errors.
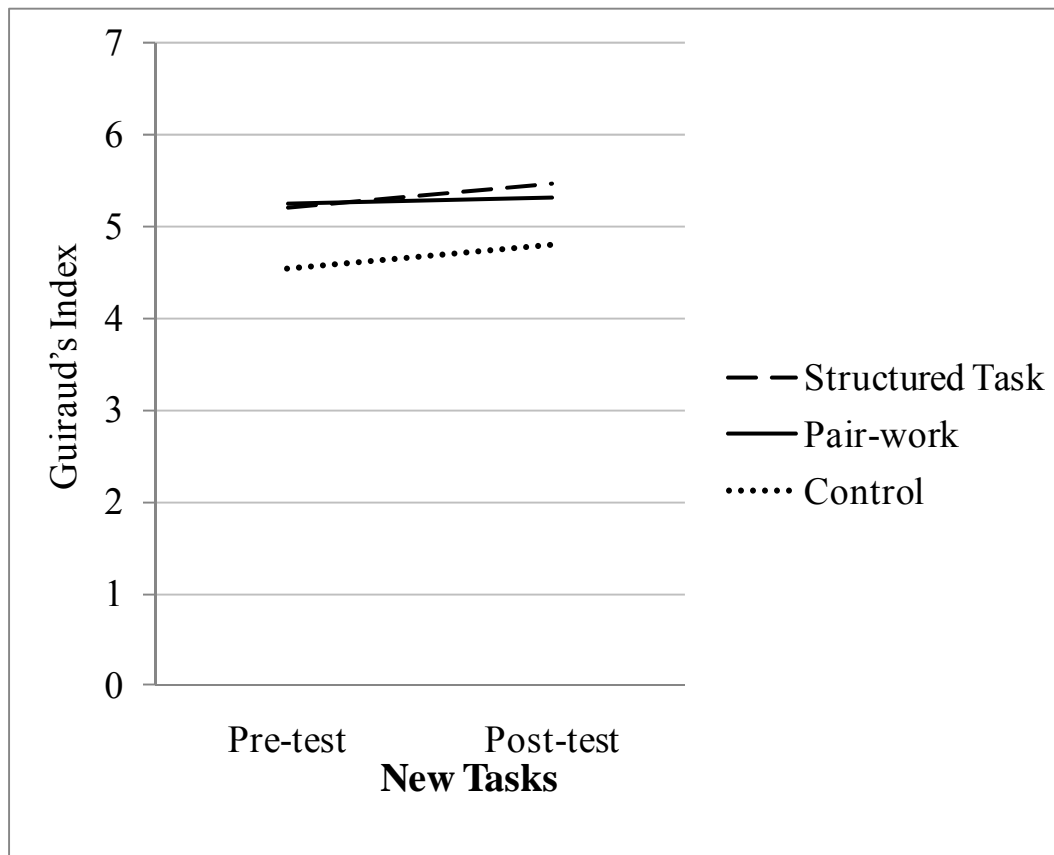
Figure 18. Guiraud's Index on the New Tasks

Since the groups differed on Guiraud's Index on the pre-test for new tasks, a one-way

ANCOVA was also conducted for this variable. The results indicate that the covariate, pre-test

scores, was significantly related to the post-test scores, $F (1, 42) = 39.764$, $p = .001$, $\eta_p^2 = .486$.

However, after the adjustment of group means on the post-test, the main effect of group was not

significant, $F(2, 42) = .642$, $p = .531$, $\eta_p{}^2 = .030$.[18] That is, the groups did not differ on gains,

which confirms the results of a mixed-design ANOVA.

4.2.5 Summary of Mixed-design ANOVA Results on New Tasks

The summary of results of mixed-design ANOVAs on eight dependent variables is

provided in Table 27. There was a significant main effect of time on pragmatic appropriateness

and vocabulary diversity. The effect sizes indicate that there was a very large improvement in

pragmatics appropriateness ($\eta_p{}^2 = .382$) and a smaller improvement in vocabulary diversity but

still a large effect ($\eta_p{}^2 = .173$). However, the interaction between time and group was not

significant for any of the variables, indicating that the groups did not differ in terms of gains. The

results also indicate that the main effect of group was significant only for vocabulary diversity.

On both pre- and post-tests, the control group's vocabulary diversity was lower than that of the

experimental groups. While there were no significant main effects of group on other variables,

the trends are the same as for research question 1: the control group was the most fluent in terms

of speech rate, but it had the lowest scores on syntactic complexity and pragmatics.[19]

---

[18] In SPSS, regression procedures are used to remove the variation in the dependent variable that is due to the covariate; then the normal analysis of variance is performed on the adjusted scores. After controlling for and removing the effect of a covariate, ANCOVA tests for differences between adjusted means.

[19] Since the control group behaved differently from the experimental groups on the pre-tests, and because there were proportionally more L1 speakers of Mandarin Chinese in the experimental groups than in the control group, but more L1 speakers of Arabic in the control group than in the experimental groups (see Table 1 in chapter 3), a follow-up analysis was conducted to examine whether this was the reason for the differences between the experimental groups and the control group. The same mixed-design ANOVA tests were conducted as the ones above on both rehearsed and new tasks, but only L1 speakers of Mandarin Chinese were included. The only different result that this analysis produced is that control group's speech rate became similar to the speech rate of the experimental groups on both rehearsed and new tasks. That is, the control group was no longer the most fluent. Thus, it appears that L1 speakers of

Table 27: Summary of Effect Sizes of Mixed-design ANOVA Results on New Tasks

Dependent Variable	Main Effect of Time	Time x Group Interaction	Main Effect of Group	Group Differences
Pragmatic Appropriateness	.382*	.109	.103	
N Syllables / Minute	.041	.019	.082	
N Pruned Syllables / Minute	.024	.016	.108	
N Pauses / Minute	.104	.031	.089	
% Pause time	.024	.006	.069	
% Error-free Clauses	.013	.037	.111	
N Clauses / AS-unit (log)	.102	.005	.128	
Guiraud's Index	.173*	.044	.292*	C < S, P

Note. The numbers represent effect sizes. S = structured task group. P = pair-work group. C = control group.
* $p < .006$

4.3 Research Question 3

What are learners' attitudes towards CASTs and pair-work activities?

Only students in the experimental groups were asked to fill out the questionnaire, and 26 out of 33 students (78%) returned the questionnaires: 17 students in the structured task group and 9 students in the pair-work group.[20] Descriptive statistics are summarized in Table 28.[21] The results are provided only for those questions that compare CASTs and role-play activities on the same aspects. The answers were considered affirmative when ratings 4 (agree) or 5 (strongly agree) were selected, neutral when rating 3 was selected, and negative when ratings 1 (strongly disagree) or 2 (disagree) were selected. In Table 28, the percentages of affirmative answers are reported.[22] Negative ratings "disagree" or "strongly disagree" were given in only four instances

---

Arabic spoke faster than L1 speakers of Mandarin Chinese. However, the relationship between groups and statistically significant results remained the same for the other variables.

[20] The students were asked to fill out the questionnaire outside of class, so the researcher did not have control over the return rate.

[21] Inferential statistics were not used due to the different questionnaire return rates between the two experimental groups.

[22] While means of Likert-scale ratings are often reported, in this study questionnaire results were reported as percentages of affirmative answers because the question regarding preferences

for the whole questionnaire dataset, and the remaining percentages of students selected the

option "neutral."

Table 28: Descriptive Statistics on the Exit Questionnaire

	CASTs	Role-Plays	Both
All Respondents (*N* = 26)			
Like	73% (19)	69% (18)	57% (15)
Prefer	50% (13)	46% (12)	4% (1)
Improves Pragmatics	92% (24)	89% (23)	85% (22)
Improves Speaking	89% (23)	85% (22)	81% (21)
Makes me Nervous	4% (1)	15% (4)	4% (1)
Structured Task Group (*N* = 17)			
Like	77% (13)	71% (12)	65% (11)
Prefer	59% (10)	35% (6)	6% (1)
Improves Pragmatics	94% (16)	82% (14)	82% (14)
Improves Speaking	88% (15)	88% (15)	82% (14)
Makes me Nervous	6% (1)	12% (2)	6% (1)
Pair-work Group (*N* = 9)			
Like	67% (6)	67% (6)	44% (4)
Prefer	33% (3)	67% (6)	0% (0)
Improves Pragmatics	89% (8)	100% (9)	89% (8)
Improves Speaking	89% (8)	78% (7)	78% (7)
Makes me Nervous	0% (0)	22% (2)	0% (0)

Note. Percentages of affirmative answers (i.e., "agree" and "strongly agree") are reported. The remaining percentage of students selected the option "neutral" in all but four instances. The number of responses is indicated in parentheses.

For the two experimental groups combined, 73% of all respondents liked CASTs, 69%

liked role-plays, and 57% liked both. Majority of the students also agreed with the statements

that CASTs and role-plays helped them improve pragmatics and speaking skills. When asked to

state their preference, 50% of respondents preferred CASTs, 46% of respondents favored role-

play activities, and 4% of respondents (1 student) did not have a preference. That is, students'

attitudes towards both types of activities were largely positive, and half of respondents preferred

CASTs, and roughly another half preferred role-plays. One noticeable difference between the

---

towards the two types of activities was not in a Likert-scale format. For comparison purposes, it
was decided to report all questionnaire data in the same way.

two types of activities is that face-to-face role-plays seem to make students slightly more nervous (or apprehensive) than CASTs (15% versus 4%).

When examining the data by group, the same patterns emerge as for the total pool of respondents. The only difference is that students tend to prefer the type of activity to which they had more exposure because more students in the structured task group favored CASTs (59%) over role-plays (35%), and more students in the pair-work group preferred role-plays (67%) over CASTs (33%). However, this is a tendency in that learners' preferences are not entirely dependent on their group assignment.

The students were asked to provide the reasons for why they preferred a particular type of activity (see Table 29 for the reasons and exact comments provided). The most common reason for favoring CASTs was the availability of native speaker input. For example, one student wrote: "I prefer to talk with native students. I can learn more good sentence and idea from them." Another reason for preferring CASTs was an opportunity for independent learning. For example, one student said: "…in contrast on role-plays, conversation depends on your partner. If your partner don't want to talk or just want to finish it fastly, you have to speak a little." Two participants also commented on the benefit of CASTs as a low risk environment. One student wrote: "By myself cannot nervous." This echoes the results described earlier that students seem to find role-plays a bit more stressful than CASTs. Finally, one person preferred CASTs because he/she found them to be more challenging than role-plays with classmates.

The most common reason provided for preferring the role-play activities was the opportunity to talk to actual people. Students thought that this is more fun or feels more real than responding to pre-recorded prompts on the computer. Several students also recognized the benefits of interaction. One participant wrote: "It helps me to interact with my classmate and my

classmate speaks English like me slowly. If I don't understand word, I tell him to say again, but

when I talk to computer, I can't listen to the word again." Other students wrote that they can

point out mistakes to each other or learn new phrases from their pair-work partners.

Table 29: Participants' Reasons for Preferring CASTs or Role-Play Activities

Reasons	Exact Comments
CASTs	
Native Speaker Access	1. "Because the person who on computer was American or good at English. I could learn how native speaker respond to others problems and solutions." 2. "I prefer to talk with native students. I can learn more good sentence and idea from them." 3. "I do it with a native speaker." 4. "I can learn something for Americans and know how American solve this problem." 5. "I can hear computer said. I can learn he or she said. I can learn more word and short sentence. This is more helpful for me." 6. "I can talk with a local speaker in computer."
Independent Learning and Native Speaker Access	7. "Online partner is American, so I can easily understand him and he speaks correctly, in contrast on role-plays, conversation depends on your partner. If your partner don't want to talk or just want to finish it fastly, you have to speak a little." 8. "Because I have my own time to speak to American person without noise."
Low Risk Environment	9. "Because I didn't see the person face to face that when I talk with anyone face to face I forget the word." 10. "By myself cannot nervous."
More Challenging	11. "I think it's harder to talk to native speaker than my classmates."[23]
Role-Plays	
Interaction with Real People	1. "It's more real." 2. "I like talk to real human." 3. "I like talk with real person." 4. "Face to computer is very boring." 5. "It's more funny. I can talk with real people." 6. "I like face to face talking. I like eye-contact." 7. "Talk with real people makes me feel more easy. And we can change something every time."

---

[23] The student probably meant to say that even though it is more challenging to talk to native speakers than to classmates, he/she prefers the use of CASTs to overcome this difficulty.

Table 29 (cont'd)

Specific Benefits of Interaction	8. "Doing role-plays in pairs, we can speak more nature and after that we can tell each other what mistake that we made."
	9. "I can see my mistakes after the role-play."
	10. "I learn new phrases if my partner have new phrases, and I can correct myself when I say a wrong grammar."
	11. "It helps me to interact with my classmate and my classmate speaks English like me slowly. If I don't understand word, I tell him to say again, but when I talk to computer, I can't listen to the word again."
	12. "Doing role-plays is more funny than talking to a computer. And talking with people can improve your English skill quickly."

Note. Each numbered comment is a participant's full answer to the question why he/she preferred one type of activity over another. Several students did not provide a reason for their preference; thus there are 23 comments rather than 26 (the number of respondents).

In summary, the majority of the students responded positively to both types of activities even though they still had a preference for one type of activity or another. It seems that students' attitudes are influenced in part by their group assignment and in part by their preference to either work independently or interact with an individual when practicing pragmatics and speaking skills.

CHAPTER 5: DISCUSSION

The results reported in the previous chapter are discussed below. For research questions 1 and 2, only quantitative analyses were conducted. However, when rating the data for pragmatic appropriateness and calculating fluency, accuracy, and complexity measures, I took a note of certain trends, which are reported qualitatively in order to add to the interpretation of findings.

5.1 Research Question 1

What is the effect of task rehearsal on improvements in pragmatics and speaking skills on the same (rehearsed) task?

This research question investigates the immediate effect of rehearsal on task performance.

5.1.1 Changes in Pragmatics and Speaking Skills Across Groups

It was predicted that the control group would improve less than the two experimental groups on pragmatics because both experimental groups, but not the control group, participated in consciousness-raising activities on pragmatics, received authentic input, and engaged in rehearsal, and the control group did not focus on pragmatics in their courses. However, none of the groups showed statistically significant improvement on pragmatics. One explanation could be that improvements were too small to be detectable through the ratings on a scale from 1 to 6. Perhaps magnitude estimation method, first introduced to SLA by Bard, Robertson, and Sorace (1996) for use with acceptability judgments and employed in Gass et al. (1999) to examine the effect of rehearsal on oral development, should be used to detect small changes. Pragmatics is also a difficult aspect of communicative competence and thus takes time to develop (Blum-Kulka & Sheffer, 1993). The evidence that participants' pragmatic competence was in a developmental stage comes from self-corrections in the data. As shown in Example 13, a participant in the structured task group self-corrects her use of the less polite phrase *I want* with a

95

newly learned more polite form *I would like*. In Example 14, a pair-work group participant self-corrects from *I thought if I can* to a very polite phrase *I wondered if I could*.

Example 13. Post-test, structured task group participant, ID 140

"Uh uh I have a question to told I have a problem to told you. Because I want to I would like to change my change my topic of this paper."

Example 14. Post-test, pair-work group participant, ID 71

"So I thought if I can if I wondered if I could stay at uh at your parents' house?"

A third possible explanation for the lack of improvement in pragmatics is that learners' attention could have been diverted to other aspects of oral performance in addition to pragmatics. (Note that during rehearsal and testing, the students were instructed to try to improve their <u>overall</u> performance, which, they were told, includes pragmatically appropriate and fluent production, correct grammar, complex sentences, and varied vocabulary). In Example 15, a participant self-corrects from a less polite form *I want to you* to a more polite form *could you help me* on a pre-test, but uses a less polite form *I want to you* on a post-test, which could be an indication that on a post-test, this participant was focusing on other issues, such as fluency, accuracy, and complexity.

Example 15. Pre-test, structured task group participant, ID 170

"So uh I want to you cou- could you help me do do my homework?"

Post-test, structured task group participant, ID 170

"Hi professor. Uh I have a big problem about about my homework. I want to uh you give me a hand."

The absence of immediate feedback and little time spent on consciousness-raising could also contribute to the lack of improvement on pragmatics. First, it appears that some students in the structured task group used input incorrectly on the post-tests. For example, in one of the CASTs, when making an argument for changing the topic of the paper, a native speaker of English, acting as a student, said: "I really really like this new topic." However, one participant in the structured task group incorporated this input into his response on the post-test as follows: "but I really really really want to change" and "so I really really want to switch." Thus, while in the input the combination *I really really like* made a positive impression on the hearer, in the L2 learner's responses the combination *I really really want* sounded very demanding in the context of a request and made this participant's post-test score lower than his pre-test score.

Input that was explained during the consciousness-raising stage was incorporated correctly by both experimental groups. One phrase that stands out is *I was wondering*. There were only two instances of this phrase on the pre-tests: one learner used it when asking a friend for help, and another learner used it when asking an instructor for help. Both learners were from the pair-work group. However, this phrase was used extensively by the experimental groups on the post-tests: there were 21 instances of this phrase in the structured task group and 17 instances in the pair-work group.[24] Nevertheless, many of the learners, even those who used a very polite phrase *I was wondering* in requests, also produced impolite or somewhat impolite phrases. In Examples 16 and 17 below, responses of one structured task group participant and one pair-work group participant are provided. In Example 16, the impolite response is "Because the first topic I think is so boring." In Example 17, the somewhat impolite response is "Actually it is just I'm not interested in uh in it, so I want to change the topic."

---

[24] Not all learners produced this phrase correctly on the post-test. For example, some variations of this phrase were *I just wondering*, *I was wonder*, and *I was wondered*.

Example 16. Post-test, structured task group participant, ID 40

"I am wondering if I could change my topic on this paper… Because the first topic I think is so boring."

Example 17. Post-test, pair-work group participant, ID 141

"I just I'm I just wondering if I can change the topic of my paper…. Actually it is just I'm not interested in uh in it, so I want to change the topic."

That is, it appears that the students learned well from positive evidence when input was first explained, but they did not have a chance to learn from negative evidence because they were not provided with ample examples of inappropriate responses and they did not receive immediate feedback. It should be also pointed out that although learners improved on pragmatics when they used phrases that were explicitly pointed out to them, such as *I was wondering if I could*, *could I*, *I would like*, etc., pragmatic ratings did not increase by much because learners still used inappropriate statements, such as *this topic is so boring*. Both raters stated that inappropriate responses often influenced their pragmatic ratings and overshadowed the use of polite phrases.

The examples above suggest that learners should either receive feedback after using CASTs for rehearsal or be introduced to specific forms, their meanings, and use before engaging in rehearsal. The importance of feedback specifically for pragmatic development has been highlighted by numerous researchers (Bardovi-Harlig & Hartford, 1993, 1996; Koike & Pearson, 2005; Martínez-Flor & Fukuya, 2005; Takimoto, 2006; Yoshimi, 2001), and feedback has been recognized as an important factor in language development in general, as several recent meta-analyses have shown (Keck, Iberri-Shea, Tracey-Ventura, & Wa-Mbaleka, 2006; Mackey & Goo, 2007; Russell & Spada, 2006). Feedback can be supplied by learners during mutual interaction, but learners at low proficiency levels may not always be able to do it well (Kim &

McDonough, 2008; Leeser, 2004; Williams, 1999). Thus, additional sources of feedback should be available. Native speaker feedback is especially important for learning pragmatics (Cohen, 2005; Sykes, 2008), but at the same time native speakers rarely provide feedback on pragmatic appropriateness because pragmatically inappropriate speaker is usually thought of as impolite rather than incorrect (S. Gass, personal communication, July 19, 2011). Feedback can, of course, be provided by an instructor, but when CASTs are used for rehearsal outside of class, it would be ideal to provide automatic feedback via a computer program. Feedback is currently not provided in the *Conversations* program, which was used to deliver CASTs in this study, because voice recognition technology is not sophisticated enough at this point to be applied to language learners (Derwing, Munro, & Carbonaro, 2000). However, some have argued that speech recognition becomes more accurate when restricted scenarios are used because learners' responses are somewhat predictable in such situations (Morton & Jack, 2005, 2010). Thus, it might be possible in the future to make CASTs adaptable to learners' responses and to provide automatic feedback. Since in this study the structured task group outperformed the pair-work group on descriptive statistics on pragmatics, it is possible that with the addition of immediate feedback, CASTs might be more advantageous than learner-learner role-plays for the development of pragmatic knowledge. This, of course, would need to be investigated.

With regard to other areas of oral performance, the results indicate that on the rehearsed tasks, there was significant improvement for all groups only on vocabulary diversity, which is surprising because previous research has shown that after rehearsing a task, learners generally improve on several aspects of oral performance. For example, in Gass et al. (1999), participants showed improvement in overall proficiency, selected morphosyntax, and lexical sophistication. In Bygate (2001), rehearsal led to improvements in fluency and syntactic complexity. Following

Skehan's (1996) proposal of the trade-off effect, a viable explanation is that due to limited cognitive capacity, language learners, especially those at lower levels, can only focus on some aspects of language production but not others even after they have engaged in planning a task. Foster and Skehan (1999) also report that based on the pooled data sets from Foster and Skehan (1996) and Skehan and Foster (1997), a selective rather than across-the-board improvement on fluency, accuracy and complexity was observed. If learners in the present study were focusing not only on the use of diverse vocabulary, but also on pragmatics, changes in which perhaps could not be detected through the rating scale used, it is possible that the participants could not focus on fluency, accuracy, and syntactic complexity at the same time. Since the tasks were communicative, learners could have been paying most of their attention to content than to structure. The same was found in the study of Németh and Kormos (2001), where in the repeated version of the task, the learners paid more attention to the informational content of the message, as evidenced by the higher number of supportive moves the learners produced.

On the other hand, these findings can be explained in terms of task complexity. The tasks that learners performed are presumed to be difficult because (1) the speech acts were of high rather than low degree of imposition and (2) learners did not have much experience with such pragmatic situations, as several of the students mentioned on the Exit Questionnaire (see section 5.4 for exact comments). In fact, learners in the experimental groups were provided with pragmatic instruction to help them with these tasks. Learners did not improve on fluency in this study, which follows the findings from previous research that planning appears to have a smaller effect on fluency in the case of more complex tasks (see R. Ellis, 2009, for a review). As in Robinson (1995), planning led to improvements in vocabulary diversity in this study. However, contrary to Robinson's findings, learners in this study did not improve on accuracy. My proposal

is that learners' attention was diverted to pragmatics even though small changes in pragmatics could not be measured immediately. Following the Cognition Hypothesis, improvements on accuracy and complexity should be on the specific aspects dictated by the demands of a complex task (Robinson, Cadierno, & Shirai, 2009). For example, when reference to motion is required, complex tasks are hypothesized to promote the use of more accurate and complex lexicalization patterns for referring to motion events. In this study, pragmatic appropriateness was required; however, judging from statistically significant findings, learners seem to have focused on more complex lexicalization patterns but not on accuracy. One prominent example can illustrate this finding. After treatment, many learners began to use a new phrase *I was wondering* on post-tests, but they also made grammatical mistakes, such as *I just wondering*, *I was wonder*, and *I was wondered.* Thus, when tasks are extremely complex or perhaps when learners are attempting to incorporate new forms and concepts, accuracy may not increase.

5.1.2 Comparison of the Three Groups

Regarding the effect of treatment, the results indicate that there were no differences between groups on gains. As mentioned earlier, there were gains only on vocabulary diversity. Surprisingly, the control group did not improve less than the other two groups on vocabulary diversity even though learners in this group did not have a chance to rehearse the tasks that they were tested on. Given that the control group started out lower than the experimental groups on the pre-test, participants in the control group could have improved as much as those in the experimental groups because they had more room for growth. The same conclusion was reached in other studies where low-performing learners improved more or as much as high-performing learners (Chiu, Liou, & Yeh, 2007; Freed, 1995; Harward, Allred, & Sudweeks, 1994; Ife, Vives Boix, & Meara, 2000). In fact, the correlation between pre-test scores and gains on vocabulary

diversity for the rehearsed task was negative (two-tailed Pearson correlation, $r = -.496$, $p < .001$), which confirms that those who scored low on pre-tests tended to have higher gains than those who scored higher on pre-tests. An ANCOVA with pre-test scores as a covariate can sometimes remedy such situations, but others believe that this use of an ANCOVA is unwarranted in non-randomized quasi-experimental designs with large group differences on the pre-test (as is the case in this study with intact classes) because it cannot be assumed that the groups have the same true baseline (Cribbie & Jamieson, 2004; Miller & Chapman, 2001; Pedhazur, 1997). [25]

While the groups did not differ on gains, there was a significant main effect of group on pausing, accuracy, syntactic complexity, and vocabulary diversity. Specifically, on the pre-test and the post-test combined, the control group had lower scores than both experimental groups on syntactic complexity and vocabulary diversity. Interestingly, the control group produced language that was more accurate than that of the structured task group (69.3% versus 57.6% of error-free clauses). Descriptive statistics indicate that the control group had the lowest scores on pragmatics. The four measures of fluency also indicate that the control group was more fluent than the pair-work group, although the difference was statistically significant only for pausing measures. (However, recall that when only participants with Mandarin Chinese as their L1 were selected, the control group's speech rate was similar to the speech rate of the experimental groups). The experimental groups only differed on the percentage of pausing time with the pair-work group pausing more than the structured task group: 6.7% versus 3.5% of total response time.

---

[25] Although in this study pre-test group differences for vocabulary diversity on practiced tasks were significant at the $\alpha = .01$ but not $\alpha = .006$ level, an ANCOVA was still conducted post-hoc; however, after pre-test differences were taken into account, the groups did not differ on the post-test.

This may indicate that the three groups, especially the control group, were focusing on different aspects of performance. Thus, the control group appears to have been paying more attention to fluency and accuracy than to pragmatics and complexity, while the experimental groups were focusing more on pragmatics and complexity. However, while individual learners may differ, for example, they may be either "accuracy-oriented (conservative erroravoiders)" or "'cutting edge' oriented (less worried about making errors, and more interested in taking chances with newly-learnt language)" (Skehan, 2002, p. 293), it is difficult to explain why there are such differences between groups because all participants were told that they need to produce responses that are pragmatically appropriate, fluent, accurate, and include a variety of different structures and lexical items.[26]

There could be several possible explanations for such group differences that existed on both pre-tests and post-tests. Most of the group differences were between the control group and either one or both of the experimental groups. First of all, it is not surprising that the control group participants were not focusing on pragmatics as they did not receive instruction and opportunities for rehearsal on this aspect. Thus, the control group participants were probably able to pay more attention to fluency and accuracy. Additionally, pragmatically appropriate language may be more syntactically complex, especially because in this study learners were taught such phrases as *I was wondering if you could* and *Could you … if you have time*. Thus, since the control group participants were not focusing on pragmatics, it is not surprising that their syntactic complexity also did not increase.

---

[26] It is possible that the control group participants did not quite understand the concept of pragmatic appropriateness since they did not receive instruction on pragmatics; however, the term *polite* was used so that the participants had at least some idea of what they were asked to do.

A subsequent more detailed examination of the data revealed that the control group produced shorter responses (measured in the total number of syllables) than the experimental groups on both pre-tests and post-tests and on both rehearsed and new tasks, and all differences were statistically significant. The descriptive statistics are presented in Table 30, and the results of one-way ANOVAs are summarized in Table 31.[27] While shorter responses could be often due to lower proficiency, it is not the case in this study because the three groups did not differ on the placement test. Rather, the control group appears to have approached the tasks very differently from the experimental groups. Perhaps because learners in the control group only took pre-tests and post-tests, but were not subjected to any treatment, they could have envisioned the tasks as being truly tests rather than opportunities for learning and improvement, and for that reason they focused highly on fluency and accuracy, producing the language they were comfortable with rather than taking risks with more complex and varied structures. This is a viable explanation because it is widely accepted that test-takers' and learners' interpretations of the task may differ from the intentions of task developers or instructors (e.g., Bachman, 2002; Chapelle, 2001; Lumley, Brown, & Zhang Wenxia, 2003).

Finally, recall that the experimental groups and the control group had different instructors, although during the study only the researcher interacted with each of the groups. This could be another factor that led to different patterns of performance between the experimental groups and the control group.

---

[27] The assumptions of normality, independence of observations, and measurement of data on interval scale were met for ANOVAs on total syllables produced. The homogeneity of variances assumption was violated for rehearsed tasks post-test and new tasks pre-test, and thus Welch's test was used to correct for this violation.

Table 30: Descriptive Statistics for Length of Responses (Total Syllables Produced)

Test		Structured Task		Pair-Work		Control	
		M	*SD*	*M*	*SD*	*M*	*SD*
Rehearsed Tasks	Pre-test	149	67	145	53	85	40
	Post-test	176	48	177	58	99	30
New Tasks	Pre-test	146	60	129	52	64	22
	Post-test	166	50	136	52	87	46

Table 31: One-Way ANOVA Results on Length of Responses (Total Syllables Produced)

Dependent Variable	*df*	*F*	*p*
Rehearsed Tasks Pre-Test	2, 47	7.295	.002
Rehearsed Tasks Post-Test	2, 28.427*	21.624	< .001
New Tasks Pre-Test	2, 24.660*	20.607	< .001
New Tasks Post-Test	2, 45	10.776	< .001

Note. For each of the dependent variables, the structured task group and the pair-work group differ significantly from the control group.
* Welch's test was used to correct for the violation of homogeneity of variances assumption.

The structured task and the pair-work groups only differed on the percentage of pausing time with the pair-work group pausing more than the structured task group: 6.7% versus 3.5% of total response time. These results should be interpreted with caution because, as mentioned earlier, the two measures involving pauses were not normally distributed, and because no other significant differences between the experimental groups were found. Additionally, while the difference in pausing time is statistically significant, conceptually the difference between 6.7% and 3.5% of pausing time may not be large. That is, this difference may not affect fluency. Moreover, longer time spent on pausing does not always result in less fluent language. Segalowitz (2010) points out that the situation is complex in terms of objective measures of fluency, including pause phenomena. Some studies have found that not only the number of pauses and pausing time, but also pause placement may affect the perception of fluency (see O. Kang, Rubin, & Pickering, 2010, for a review). Thus, mid-clause pauses but not end-clause pauses may lead to lower fluency (Pawley & Syder, 2000; Tavakoli & Foster, 2008).

5.2 Research Question 2

To what extent can task rehearsal influence performance on a new task?

While the immediate effect of rehearsal was investigated through research question 1, the effect of rehearsal on new tasks (i.e., whether task rehearsal facilitates pragmatic and oral development in general) can be observed through research question 2.

5.2.1 Changes in Pragmatics and Speaking Skills Across Groups

The results indicate that for new tasks, all groups improved on pragmatic appropriateness and vocabulary diversity, and the effect sizes were large ($\eta_p^2 = .382$ and $\eta_p^2 = .173$ respectively). This is one piece of evidence in support of the earlier statement that learners were paying most attention to content than to structure or form when performing the tasks in this study. The fact that students produced longer responses (in the total number of syllables) on the post-test than on the pre-test (see Table 30) may also indicate that learners were focusing on providing more informational content rather than on polishing the structure.

While no significant changes on pragmatics were observed or could be measured via the rating scale used on immediate (rehearsed task) post-tests, delayed effects of planning on pragmatic development were observed on new task post-tests.[28] Delayed effects have been proposed (Gass, 1997; Mackey & Philp, 1998) and found (Mackey, 1999; Mackey & Goo, 2007) to be stronger than immediate effects for advanced structures. Thus, the results of this study suggest that pragmatic knowledge takes time to develop and should be measured not only immediately after treatment, but also after some period of time.

5.2.2 Comparison of the Three Groups

---

[28] Recall that post-tests on new tasks were conducted at the end of the study, while rehearsed task post-tests were conducted once a week immediately after a task was rehearsed.

The results also indicate that for new tasks the main effect of group was significant on vocabulary diversity. Specifically, on pre- and post-tests combined, the control group's vocabulary diversity was lower than that of the experimental groups, which is the same as the result on the rehearsed tasks. There was no significant main effect of group on other measures for new tasks, but trends (i.e., descriptive statistics) were the same as those for rehearsed tasks: the control group spoke faster and produced less complex and pragmatically appropriate responses than the other two groups. Thus the same explanations as before could apply: (1) the control group focused less on pragmatics than on the other aspects because it did not receive instruction in this area; (2) it viewed the tasks as a test rather than a learning opportunity and thus focused on accuracy and fluency; (3) different instructors for the experimental groups and the control group could have led to different patterns of performance between groups.

However, the main goal of this study was to examine the effect of treatment: rehearsal of a task via CASTs, via role-plays in pairs, or lack of rehearsal. No significant differences in gains were found between the groups on either rehearsed or new tasks. If there are differences between the two experimental conditions, the use of a more adaptive computer program that provides feedback might be necessary to uncover them. In this study, it might be the case that the advantages and disadvantages of CASTs and learner-learner interaction cancel each other out, equalizing the two conditions. The advantage of CASTs is that they provide native speaker input, which is not available in learner-learner interactions. On the other hand, the advantage of learner-learner interactions is that they are natural and allow for the provision of feedback, while in CASTs the prompts are pre-planned/non-adaptive and feedback is not available. Although teacher-guided planning has been found to be more beneficial than learner-learner planning in previous research (Foster & Skehan, 1999), rehearsal via structured tasks might outweigh the

benefits of rehearsal in dyads only when it resembles teacher-guided planning more closely, that is, when interactive feedback is also provided.

Regarding the effect of rehearsal, it was predicted that on new tasks the experimental groups would outperform the control group because the effect of rehearsal would transfer to new tasks. However, on vocabulary diversity all three groups had similar gains. Since the control group started out lower than the experimental groups on vocabulary diversity on the pre-test, and this difference was significant for new tasks ($p < .001$), participants in the control group could have gained as much as those in the experimental groups because they had more room for growth. That is, while for the experimental groups mean Guiraud's Index was above 5 on both pre-tests and post-tests (for structured task group, it changed from $M = 5.21$ to $M = 5.46$, and for the pair-work, it changed from $M = 5.25$ to $M = 5.31$), for the control group, mean Guiraud's Index was below 5 ($M = 4.54$ on the pre-test and $M = 4.79$ on the post-test). In fact, the correlation between pre-test scores and gains on vocabulary diversity for the new task was negative (two-tailed Pearson correlation, $r = -.292$, $p = .044$), which confirms that those who scored low on pre-tests tended to have higher gains than those who scored higher on pre-tests.

It was also unexpected that the control group would gain as much as the experimental groups on pragmatics because little attention was paid to pragmatics in the students' regular language courses. However, it is important to point out that although the control group improved on pragmatics, it still received the lowest ratings on both the pre-test and the post-test. That is, the control group is in some way different from the experimental groups even though their gains are the same. Since students were in the target language environment, the control group could have encountered the same tasks in real life that the experimental groups rehearsed in class, such as asking a friend to look over their paper, asking an instructor for permission to change the topic

of the paper, etc. Several studies have shown that pragmatic development can be enhanced through exposure to the target language environment (e.g., Matsumura, 2001; Shardakova, 2005). Thus, studies conducted in a foreign language environment might be necessary to partial out the influence of target language environment.

However, while the control group's exposure to the target language might explain why there were some gains on pragmatics for this group, this factor should not be enough to equalize the gains of the experimental groups and the control group because explicit instruction on pragmatics is also necessary. Thus, another reason that the experimental groups gained as much as the control group on pragmatics could be that the instruction which the experimental groups received was not explicit enough (i.e., feedback was not provided as part of the experimental treatment). Additionally, given that pragmatic development takes time, pragmatic instruction provided only once a week for three weeks (or six weeks if apology speech acts are counted) might have not made a noticeable difference in pragmatic development of the low-level learners in this study.

In summary, the differences between the experimental groups and the control group on pragmatics could have been reduced by two factors: on the one hand, the control group likely had unanticipated exposure to pragmatics in their L2 environment, and on the other hand, the experimental groups might have not received enough instruction and feedback to differentiate them from the control group.

While previous research findings are that rehearsal generally improves subsequent performance on the same (rehearsed) tasks (Bygate, 2001; Bygate & Samuda, 2005; Gass et al., 1999; Lynch & Maclean, 2000, 2001; Németh & Kormos, 2001), but not necessarily on the new tasks (Bygate, 2001; Gass et al.), in this study, the influence of rehearsal on same and new task

performance is difficult to assess due to the kind of pragmatic instruction provided and because the no-rehearsal group (i.e., the control group) differed from the experimental groups on pre-tests.

5.3 Discourse Patterns of the Experimental Groups during Rehearsal

Even though the two experimental groups did not differ on gains, as evidenced by the quantitative analyses above, discourse patterns of the structured task group and the pair-work group on one task were compared to examine whether there were any qualitative differences between these groups.

Task 2 from the rehearsal stage was examined. The situation for task 2 was:

Role X

You are a student. Ask your instructor if you can change the topic of your paper. The paper is due soon. You realize (= you know) that you do not have a lot of time to work on the new topic. However, you have good reasons for changing the topic of your paper. Try to get permission from your instructor to change the topic. *Be polite*.

Role Y

Imagine that you are an instructor. Your student asks you if he or she can change the topic of the paper. But you think that it is too late to change the topic now because the paper is due soon. Try to convince the student to write the paper on the original topic. *Be strict but polite*.

The data from all participants in the structured task group and in the pair-work group during the rehearsal of task 2 were examined. Common trends within each group were identified, after which the responses of several participants from each group that exemplified these trends were selected.

First, rehearsal patterns for three participants from the structured task group are provided. Note that the participants rehearsed a task several times, and all iterations of the task are provided in order to show progression in language development for each of the three participants. Also note that the learners were given 15 minutes to rehearse a task as many times as they could, and thus the number of times that each participant was able to rehearse a task differed. The number of rehearsals for each student ranged from one to seven, but most students rehearsed a task two, three, or four times. In the examples that follow, participant 40 was able to rehearse task 2 six times, although the first, third, and the last iteration were not finished; participants 200 and 140 were able to rehearse this task three times. It should be also pointed out that participants rehearsed both roles in each task. For example, in task 2 they rehearsed the roles of a student requesting permission to change the topic of the paper and the role of an instructor. However, in the examples from the structured task group provided below, only the role of the student is provided because the goal of the study was to examine how participants learn the speech acts of requests and refusals (specifically, declining a suggestion), but not how they learn to talk as instructors.

When switching roles in CASTs, participants received a model of a response they could provide as a student, which is one kind of input that the structured task group received. The model for a student role in CAST 2 is provided in Table 32 below, and it will become clear in the examples that follow that participants incorporated this input into their responses.

Input consisted of CAST prompts from both roles: the instructor prompts provided in the examples below and the student role prompts listed in Table 32. Students would start their rehearsal by answering the prompts from an instructor, and thus the input they first received was from an instructor. Note that in the transcripts below, the phrases and ideas from input are

Table 32: Model Responses for a Student Role in CAST 2

Prompt Number	Model Response
1	Hi. I was wondering if I could switch topics for my paper. I know it's due soon. But I'm not finding very much stuff on this topic, and there's another topic that I can find a lot of information on.
2	I'm really really having problems with my first topic, and I've already found a lot of stuff on the second topic, and you know like if I can handle switching topics, would you allow it?
3	I understand, but I really really like this new topic. So what if I come up with an outline and get it to you tomorrow?
4	Yeah, I'll think about it some more. And when I make a final decision, I'll let you know. Thanks.

underlined, and words/phrases that became correct after a few iterations of the task are bolded.

There are two possibilities for each of the words, phrases, and ideas from input that the

participants were incorporating into their responses: (1) they either already knew the word/phrase

or they knew that making a particular argument could lead to a successful and polite request, in

which case the participants were primed to use that word/phrase/argument due to input; or (2) the

participants learned the word/phrase/idea from the input in CASTs. In fact, some of the phrases

that appear in CAST prompts, specifically *I was wondering if I could* and *What if I*, were

introduced to students in class. In either case, there is a general trend that participants are

incorporating phrases and ideas from the input in CASTs because they did not use these phrases /

ideas during the first iteration of the task.

I will now look at the concrete examples. In Example 18, participant 40 did not finish the

CAST during the first rehearsal. During the second iteration, some phrases from input already

seem to be appearing: *I was wondering* and *I will let me know*, although the phrase *I will let me

know* is not entirely correct. Participant 40 did not finish the third iteration of the task. During the

fourth iteration, this participant is beginning to use the ideas from input: she found more

information on the new topic and she really likes the new topic. That is, while she still makes

negative statements, such as "I really don't like the second topic" (most likely meaning the first/original topic), there are now more positive statements, which are more effective arguments. This participant's phrase *I will let me know* is still not entirely correct in the fourth iteration. During the fifth iteration, participant 40 is using more words and phrases from input, such as *switch* instead of *change*, *I know it's due soon*, *stuff* instead of *information*, and *I understand* instead of *I know*. This participant also finally uses the phrase *I will let you know* correctly. Participant 40 does not finish the sixth rehearsal, possibly because the 15-minute rehearsal period was over.

> *Example 18. Responses of participant 40, acting as a student, to prompts in CAST 2*
>
> *Iteration 1*
>
> Prompt: How can I help you?
>
> Participant: Hi professor. Uh I really have some problem because uh I want to change the topic of my paper. [pause] Because I don't like this topic and I think this topic is same uh is same as last one, so I want to change it.
>
>   [This participant did not finish the first rehearsal of CAST 2.]
>
> *Iteration 2*
>
> Prompt: How can I help you?
>
> Participant: Hi Missis White. I want to change my topic of my paper because I don't like it. Uh this topic is so boring and [pause] it's the same as the last one, so I really want to change it.
>
> Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

113

Participant: Um, but I re- really don't want don't want to do the this topic and I really wondering to change the topic. Because um the last topic I I get more effect and I want to change the new one.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Yeah, I know. Um. But um could you do something could you uh do something effort because I really don't like it. Um I think I will be a really good job on this on new topic and um I <u>I was wondering</u> I can change it.

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Okay, thank you a lot. I think I will be uh do my best in this topic.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, Um. **<u>I will let me know</u>** if I decide to change it.

Iteration 3

Prompt: How can I help you?

Participant: Hi uh Mister White uh Missis White oh my God.

                 [This participant did not finish the third rehearsal of CAST 2.]

Iteration 4

Prompt: How can I help you?

Participant: Hi uh I <u>was wondering uh if I</u> can change the [pause] topic of my paper because uh <u>I can find um I CAN FIND more information on my new topic</u> and I really don't like the second topic.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: Yeah, but <u>I really really like the se- uh new topic</u> and I really really want to change it. Because the first topic I don't uh I do it very bad. <u>I don't found any information on it</u>. <u>But I can found another information on the second topic, so I really like the new topic</u>.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Yeah, I know, um, but <u>I really like this topic and I will uh outline to put it</u> so uh

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Okay, so it's very important to change this so if I want to change it **<u>I will let me know</u>**. So thanks a lot.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, thanks a lot. Bye.

Iteration 5

Prompt: How can I help you?

Participant: Hi uh <u>I wondering</u> [pause] <u>if I could chan- switch my topic</u> of my paper. <u>I know it's due soon</u>, but <u>I really like the new topic</u>.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: Yeah, I <u>REALLY REALLY like my new paper my new topic</u>. Because uh if I change the new topic, <u>I will find more information and more</u> [pause] <u>stuff</u> on my new topic. So I really really want to change it.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: <u>Yeah I understand</u>. <u>I will put it in the uh [pause] outline</u> and and I will do my best on this topic, so trust me.

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Uh it's okay. Uh, I will do my best on this paper, and if I want to if I decide it, **<u>I will let you know</u>**. So thank you.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, uh sure. Uh thanks very much.

Iteration 6

Prompt: How can I help you?

Participant: Hi <u>I was wondering if I could if I I was wondering I WAS WONDERING if I can switch</u> the new to- uh switch uh switch the topic of my paper because uh because I really like my new topic.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: I really really want to my new topic because uh my first topic works very bad. I I'm <u>already found more stuff</u> on my new topic and I really really want to change it.

[This participant did not finish the sixth rehearsal of CAST 2.]

In Example 19, responses of participant 200 are provided. Just like in Example 18, the participant gradually starts to incorporate ideas and phrases from the input. During the first rehearsal, participant 200 does not use many ideas and phrases from input. His utterance "I was wonder if if you have if you have some free time" during the first iteration is not based on the input from the student model because by this time participant 200 has only heard the instructor prompts. Instead, the use of this phrase could have been prompted by the previous explanations in class. This participant is trying to use one new phrase from input, although incorrectly. He heard the instructor's response "just let me know what you decide" and then replied "Okay, I will send you top- send you a email to make sure about my uh decide [pause] decide." That is, it

117

seems that this participant is trying to figure out how to use the noun *decision* and the verb *decide*. During the second rehearsal, participant 200 begins to use many ideas from the student model: he is having trouble with the original topic, he has found a lot of information on the new topic, he knows that the paper is due soon, and he inquires if he can send a new outline to the instructor. However, at the same time the phrases used to express these ideas are not target-like. For example, this participant uses such phrases as *I know it will become soon* instead of *I know it is due soon*, *I really meet a big trouble* instead of *I am having problems*, and *I was wonder could I* instead of *I was wondering if I could*. One target-like phrase from the input that this participant begins to use is *I will let you know*. When rehearsing this CAST the third time, participant 200 continues to use ideas from the student model, and there is also some evidence that the phrases are becoming more target-like. Thus, during the third iteration this participant says "I know that due day will be soon," and he also seems to understand better the difference in the use of *decide* and *decision*: "if I make new decide if I make sure the decision." However, other phrases are still not target-like: *I'm really meet a big problem* and *I was wonder could I change it?* This participant also incorporates the word *seriously*. The instructor says "I would seriously consider," and participant 200 replies "I will seriously think ab- I will think about this question seriously."

> *Example 19. Responses of participant 200, acting as a student, to prompts in CAST 2*
>
> *Iteration 1*
>
> Prompt: How can I help you?
>
> Participant: [pause] Hi professor. Uh I want to change my topic about our assessment.
>
> Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I

118

really do think you should stay with your original topic.

Participant: Yeah, but uh you know I already got a [pause] I already brainstorm and about my new topic and I already have the um outline. I think I can do it soon, and it's not waste a lot of time. So

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Oh [pause] but I think I still have I I can do it soon. It don't it doesn't cost a lot of time.

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Okay, okay, um I was wonder if if you have if you have some free time, could I send you a email to talk about um to ask you a few question about my new topic if I may.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, I will send you top- send you a email to make sure about my uh **decide** [pause] **decide**.

Iteration 2

Prompt: How can I help you?

Participant: Hi professor. <u>I was wonder could I</u> change the topic for my assessment?

Because there I meet <u>I really meet a big trouble</u>. <u>There isn't a lot of</u>
<u>information I I can find from the old topic the original topic from the</u>
<u>promotion pro- promised topic</u>. So I was wonder could I change the topic.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a
               few weeks ago, maybe we could have worked something out. But uh I
               really do think you should stay with your original topic.

Participant: Yeah, I totally understand. I know it [clears throat] Sorry, I totally
               understand. <u>I know it will become soon</u>. But I I already made the plan, so
               I don't think it will cost me a lot of time.

Prompt: Yeah, I understand. But my chief concern is that we've already met about
               the previous topic and I've looked at outlines for the class. So if you did
               change your topic now, you'd be handing in a paper without any of that
               previous guidance.

Participant: Oh yeah, really, but I really want to do it. <u>How about I send you my</u>
               <u>outline first</u>? Then could I uh could I could I change my topic?

Prompt: Okay, so you can change your topic, but I think you are taking a big risk
               because I'm not going to have much time to help you with it since we are
               approaching the end of the semester. In fact, if I were you, I would
               seriously consider whether you wanna do this or not.

Participant: Oh, but [pause] but <u>I can find a lot of information from this topic</u>. Um
               could I write could I send you a email about the new topic? Uh then if I
               have enough time, could you help me to figure this out?

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay. Uh if I decided, <u>I will let you know as soon as possible</u>. Thank you.

Iteration 3

Prompt: How can I help you?

Participant: Uh yes, professor. I I would like to change my I would like to change my topic from the assessment. Because <u>I'm really meet a big problem</u>. I <u>I cannot find a lot of information from this topic</u>. So I was won- <u>I was wonder could I</u> change it?

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: Uh yes, I know that I <u>I know that due day will be soon</u>. But <u>I really I really like the new topic</u>. And I'm really interest in this topic. <u>And I cannot find I cannot find much as much information from the original topic</u>. So could I change it?

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Yeah, but <u>how about I can send you the outline</u> and some brainstorm to your email, then you can get some information from it. [pause] Then you can give me more advice?

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are

approaching the end of the semester. In fact, if I were you, I would

seriously consider whether you wanna do this or not.

Participant: Okay, I will <u>seriously</u> think ab- I will think about this question <u>seriously</u>.

Uh so if I got any <u>if I make new **decide** if I make sure the **decision**</u>, I will

send you a email or call you to <u>let you know</u> as soon as possible.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, okay, I will <u>let you know</u> as soon as possible. Thank you.

The last example from the structured task group is Example 20. Just like in the previous

two examples, participant 140 begins to use ideas and phrases from the input during the second

and third rehearsals of a task. (Note that this participant's statement during the first iteration that

she is interested in the new topic is not from the input because by this time she has only received

input from instructor prompts.) During the second rehearsal, this participant uses several ideas

from input: she has problems with the original topic, she cannot find enough information on the

original topic, and she has more information on the new topic. However, she also expands on

these ideas from input and formulates them in a different way by saying "I can't uh find the

research from the internet from and from the library" and "if I use a new topic, I can have a lot of

support point." This participant also uses the phrases *I understand* and *make a decision*, although

she is experimenting with the forms *I understand* and *I understand it*. During the third rehearsal,

participant 140 formulates the ideas used previously in a different way: "I can't find a lot of

information to support my old pap- uh old topic" and "I really really have problem with this uh

old topic" (the latter is very close to the input "I'm really really having problems with my first

topic"). This participant also begins to use the words *switch* and *stuff* and attempts to use the

122

expression *the paper is due soon*: "I know the due day the uh the paper is due soon." This participant seems to be settling on the non-target-like version *I understand it*.

Example 20. Responses of participant 140, acting as a student, to prompts in CAST 2

Iteration 1

Prompt: How can I help you?

Participant: Hi professor. Uh I uh I would like to told you I want to chan- change the topic for my paper. Can I can I do that?

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: I know it, um, but uh I can't I can't uh write I wri- write the deeply uh deeper write on on this paper. I just know a little about that. But uh however my new my new topic I I'm very interested in this and I can I can wrote a lot, and I think it's help my my score.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Yes, maybe maybe uh I'm very appreciate. So I [pause] okay I still write this papers. Tomorrow I give you the whole paper.

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would

seriously consider whether you wanna do this or not.

Participant: Okay, I I will think about it. And uh if I uh I still still use this old topic to do it. Thank you very much.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, uh s- see you later.

Iteration 2

Prompt: How can I help you?

Participant: Yeah professor. Uh I really really want to change my topic because <u>in this old topic I have big problem to finish</u>. <u>I can't uh find the research from the internet from and from the library</u>. And I <u>I think I if I use a new topic, I can have a lot of support point</u>. And the logical is very clearly. I think it's help my it's very ea- easy to finish it.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: Yeah, I <u>I understand</u>. But <u>the old topic I have a lot of problem in it</u>. And uh yeah, I could try but I don't think I can have a uh I am very good at it um I'm not very good at it. So uh if I use this paper, uh maybe I only wrote a short uh short essay and cannot respect my own opinion in this paper.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that

previous guidance.

Participant: Okay I uh as soon as possible finish this paper and can we can I can I bring this paper tomorrow to let you loo- have a look it?

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Yeah, I understand it, you are very busy. And oh I will I will um to finish and tomorrow and I will told you my my uh make a decision and I will talk to told you that.

Prompt: Okay, either way, just let me know what you decide.

Participant: Oh, uh I'll call you after I make a decision.

Iteration 3

Prompt: How can I help you?

Participant: Yes professor. I nee- I I I would like to switch my topic because I can't find a lot of information to support my old pap- uh old topic. And I don't thi- I think my paper maybe cannot very good if I use the old old topic so I choose the new topic. I know the I know the due day the uh the paper is due soon. But uh I but I if I use the old topic, I cannot finish before the due do- due day.

Prompt: Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.

Participant: Oh yeah <u>but I really really have problem with</u> this uh old topic. I found a lot of people to help me, but I still cannot write uh enough thing. And do you have to time to help me? And [pause] I can do it better maybe.

Prompt: Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.

Participant: Yes maybe I can change, may- maybe it's right. But uh I need to change it. And the new topic, I found a lot of <u>stuff</u> to help me. Do you think it's a good good decision?

Prompt: Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.

Participant: Okay, <u>I understand it</u>. Uh okay, if I <u>make a decision</u> to do it or not change it, I'll told you as soon as possible.

Prompt: Okay, either way, just let me know what you decide.

Participant: Okay, uh uh thank you very much.

The general observation is that in the structured task group participants mainly began incorporating input into their subsequent rehearsals. Specifically, they incorporated new words and phrases, as well as ideas, such as arguments for changing the topic of the paper. Some students were also able to make certain phrases more target-like after rehearsing a task several times. That is, after several rehearsals the learners seem to be able to focus on forms. However,

learners do not always incorporate phrases from input correctly; they may analyze them in their own way. For example, participant 200 uses *I was wonder* all the time even though the input is *I was wondering*. Even when learners incorporate words or phrases from input correctly, they do not blindly repeat the phrase as a whole chunk. For example, participant 200 said "I will think about this question seriously" after he heard "I would seriously consider whether you wanna do this or not."

The examples provided above are representative of the structured task group. That is, while there were some cases with somewhat different patterns (for example, one student previewed the prompts for task 2 first without responding to them), these cases are not provided as examples because they are rare and thus do not represent the group as a whole.

As can be seen from the examples above, CAST prompts and participants' responses sometimes did not match in terms of sequencing. For example, in iteration 2, participant 200 said "Oh, but [pause] but I can find a lot of information from this topic. Um could I write could I send you a email about the new topic? Uh then if I have enough time, could you help me to figure this out?" After this, the participant heard the prompt "Okay, either way, just let me know what you decide." That is, the participant did not receive an answer to his question. However, the common pattern is that once the students know the sequencing of the prompts after several rehearsals, they begin to modify their responses, if necessary, so that the prompts and the responses match better. Thus, in iteration 3 participant 200 does not ask anymore whether the instructor can look at his email, but instead says: "Okay, I will seriously think ab- I will think about this question seriously. Uh so if I got any if I make new decide if I make sure the decision, I will send you a email or call you to let you know as soon as possible." While this design is not ideal, and the tasks would seem more naturalistic if prompts in CASTs adapted to learners' responses, CASTs are still

beneficial because they provide learners with native speaker input, which learners subsequently

incorporate. It also seems that structured rehearsal allows learners to focus their attention on

forms.

Next, rehearsal patterns for four participants (i.e., two dyads) from the pair-work group

are provided. Again, all iterations of the task are provided in order to show the development for

each of the participants. Just like in the structured task group, learners in the pair-work were

given 15 minutes to rehearse a task as many times as they could, and transcripts of the full period

of interaction are provided. An interesting finding is that learners in the pair-work group were

able to rehearse a task fewer times than those in the structured task group because role-plays

were longer than the responses learners provided to prompts in CASTs. Some learners in the

pair-work group were able to rehearse each role of a task two (or rarely three) times, and other

pair-work participants were able to rehearse each role only once. Another reason for fewer

rehearsals in the pair-work group is that some dyads also engaged in discussion (rather than

rehearsal) part of the time. Because students were sometimes not sure what they could do

differently during the second or the third rehearsal, they were told that they could give each other

advice on pragmatics, vocabulary, accuracy, complexity, and content. Most dyads chose to

include some discussion in addition to rehearsal. The most common point of discussion was

content: how students and instructors should behave in these role-plays (e.g., whether the

instructor should or should not let the student change the topic of the paper) or what arguments a

student could provide for changing the topic of the paper. This is not surprising given that role-

play situations were rather general. Students rarely discussed form. To give students more ideas

for rehearsal, they were also encouraged to use what they learned on pragmatics from the

consciousness-raising phase. Another suggestion given to students was to modify some aspect of

a role-play for each new rehearsal but to still follow the scenario provided. For example, during

each rehearsal, they could provide a different reason for changing the topic of their paper.

The transcript of interaction in the first dyad is provided in Example 21. (Note that

pseudonyms are used). In this dyad, the students role-played once, then switched roles and role-

played again, then began a discussion of the task, and then role-played one more time. First, I

examine participants' discussion (lines 53 to 91). One participant (Sara) makes a brief mention

of the polite phrases that a student can use, such as *sorry* and *I wonder*. After that, the discussion

focuses on the behavior of a student and an instructor. The other participant (Jenny) is having

difficulty being strict but polite as an instructor. She says "Uh as a teacher I think um you should

be kindly but strict. It is hard for me to do this. I don't know how to refuse the student's uh

student's question." Note that during the first rehearsal, when acting as instructors, both

participants in the end let the student change the topic of the paper. However, during discussion

Sara and Jenny disagree on whether the student should be allowed to change the topic. This

negotiation of the task might be taking place because according to the scenario the student and

the instructor should be persuading each other (politely), but it is not specified who "wins" in the

end. Sara and Jenny are not able to reach agreement, and they decide to role-play again (see lines

81 to 83). This time both participants become less polite as the role-play progresses (lines 100 to

117): Sara tries to be a strict instructor and to not let Jenny change the topic, while Jenny

strongly believes that she should be able to change the topic. Both participants first try to reason

with each other (lines 100 to 109): Sara says that the topic cannot be changed because all

students have the same topic, while Jenny explains that she tried to work on this topic first, but

she found the topic to not be "suitable" for her. Sara still tries to convince Jenny to stay with the

same topic, but offers help. After this both the student and the instructor become somewhat

129

impolite (lines 110 to 117). Jenny, as a student, says "No no I don't think the topic is very difficult. I just I'm just not interested in it." After which Sara, as an instructor, replies "But I think uh the homework is not what you are interested in, it's required. So this is your homework, so you must do it." After this Jenny becomes even more inappropriate because her response looks more like a demand rather than a request "…so I want to change another. I think it is available. I think there is so many available topics for me to choose." Participants then talk in their L1, perhaps trying to decide how to finish the role-play, after which Sara lets Jenny change the topic (unwillingly, judging by the intonation).

Thus, Example 21 reveals one difference between CASTs and role-plays in pairs: students might be more pragmatically appropriate when responding to instructor prompts in CASTs because these prompts resemble native-speaking instructors more than language learners do in role-plays. During role-plays, participants acting as students appear to be less appropriate if their partners do not act like native-speaking instructors. (This might be the reason why the structured task group is the only one that improved on pragmatics on rehearsed tasks, although this difference was not statistically significant). For example, in CAST 2 the instructor prompt was indirect and in a form of a strong suggestion rather than prohibition: "In fact, if I were you, I would seriously consider whether you wanna do this or not," while in Example 21 Sara's response, in the instructor role, was direct and strictly prohibitive: "So this is your homework, so you must do it." Of course, we should expect that language learners at the age of traditional college students (the mean age was 20 in this study) would not know how to talk as native-speaking instructors, which is where CASTs become particularly useful.[29] Whether the learners

---

[29] When role-plays involve peers, such as when a friend needs to ask another friend for help, input from CASTs might be less crucial if language learners are better equipped to act as

are able to imagine that they are speaking to an instructor (or any other authority) may depend not only on the language used (e.g., indirect but strict versus direct), but also on the accent and the appearance of the person. (It may be beneficial to investigate in the future what affects learners' perceptions).

Regarding the use of input, when role-playing, participants relied on any sources available to them: the language used in the description of the situation (e.g., Sara, lines 42-43: "But you know the due the the paper is due soon"), expressions introduced in class in previous lessons (e.g., Jenny, line 31-32: "Can I talk you- with you very quick?"), and language from the video used in consciousness-raising phase that day (Request Video 2, Appendix D) (e.g., Jenny, line 92: "Uh, excuse me. Are you professor [pause] Adams?" Sara, line 97: "Nice to meet you to. [pause] You can have a seat," and Jenny, line 98: "I was wondering I am really interested in…"). However, what is not available to participants in the pair-work group, as opposed to the structured task group, is models of how they can respond as students or instructors in the specific situation that they are rehearsing.

Example 21. Transcript of interaction between participants 41 and 141 when rehearsing task 2

1 Sara: Hi Jenny

2 Jenny: Hello, Sara. How are you?

3 Sara: Fine, and you?

4 Jenny: Uh, pretty good.

5 Sara: Okay, uh I want to ask you something about the topic.

---

friends/classmates than as instructors, but input from CASTs in such situations should still be valuable given that it would be authentic input.

6   Jenny: Topic? Okay, sure, go ahead.

7   Sara: Uh um I think this topic for me it's not very well, so I want to change the

8   topic. Can I do that?

9   Jenny: Uh, wait a moment. You want to change a the topic?

10   Sara: Yes

11   Jenny: But you know the due the the paper is due soon.

12   Sara: But I think this topic for me is a little difficult, I can't do it very well. So I

13   want to change the topic. I think if I have a new topic, I can do it well.

14   Jenny: Uh why not why don't you change it uh soo- uh before? 'Cause it's due due

15   soon.

16   Sara: Uh

17   Jenny: I think you should change uh before the due time uh before the day or it's

18   much better for you to uh finish the topic very well.

19   Sara: Because um this days I'm very busy uh so today I see the topic I think it's

20   little uh it's not fit for me so I want to change it.

21   Jenny: Uh so you can promise you shou- uh you can uh finish it on time?

22   Sara: Yes, of course.

23   Jenny: Um if you stick to change your topic, uh I'm okay, you can change. And

24   make sure you can hand in it on time and you can do it pretty perfect.

25   Sara: Okay, thank you very much.

26   Jenny: Uh you are welcome. Okay change.

             [Students switch roles here.]

27   Jenny: May I come in?

28 Sara: Okay, come in.

29 Jenny: Uh, how are you, Sara?

30 Sara: Uh, fine. What about you?

31 Jenny: Uh good. Uh, I have a question for you. Can I talk you- with you very

32   quick?

33 Sara: Okay, of course.

34 Jenny: Um okay. Um I want to change my topic of my paper. Can I do that?

35 Sara: Oh why you want to change your topic?

36 Jenny: Uh 'cause it is it is not suitable for me to for the study, you know.

37 Sara: Why? I think uh a lot students think this topic is good.

38 Jenny: Yes, I tried to do at first I tried. But I gradually I found it is really not

39   suitable for me to to further study, and I'm not sure I can finish it very well.

40   So I want to change my topic. Uh if I want if I can change my topic, I am

41   much uh I'm interested in it, so I can do it very well.

42 Sara: Okay. Um but I think it's too late to change the topic now because the

43   paper is due soon.

44 Jenny: Oh, really? [pause] However, I have uh I can do it 'cause I'm interested in

45   the uh topic so I can uh change my schedule and move so much more time

46   on it. So I think it's okay. I will hand in it on time.

47 Sara: Uh can you promise this thing? You can do it before the due time?

48 Jenny: Sure I can, I promise.

49 Sara: Okay if you can do it uh very quick uh very soon, you can change the

50   topic. And I think you can do it very well.

51	Jenny:	Oh, that's very kind of you. Thank you.
52	Sara:	You are welcome.
		[Students begin a discussion here.]
53	Jenny:	Just just five minutes we talk about it. We can talk about some advice.
54		What's your opinion how to be polite but strict by teacher?
55	Sara:	Um polite I think I can say some more such as sorry or I wonder or others.
56	Jenny:	Uh as a teacher I think um you should be kindly but strict. It is hard for me
57		to do this. I don't know how to refuse the student's uh student's question.
58		Like I don't know how to redu- re- refuse so it's hard I always agree with
59		you and then promise if you want hand in it on time, you can change your
60		topic.
61	Sara:	I think as a teacher if your student tell you this topic it's difficult, you don't
62		just ask ask him her why it's difficult, you can ask them where it's difficult
63		and you can help her to uh to sov- to
64	Jenny:	sovolve?
65	Sara:	sovolve this problem.
66	Jenny:	But I think uh if you help the problem with the students, she maybe not
67		want to change the topic. But the students' goal is to change the topic. Uh
68		as a student you should tell tell the teacher the uh reason why you want to
69		change you topic
70	Sara:	But I think if any student change the topic, other student maybe think it's
71		not good. So as a teacher we must do our best to change the student mind.
72		Um

73   Jenny: Yes but it is hard for for us to solve this problem. How you can refuse the

74   students' advice. And polite but strict. So it is not easy for me to refuse.

75   Finally I will agree with the student. And okay if you hand it in on time, I I

76   I think you can change your topic.

77   Sara: But I think as a student, although the topic is very very difficult, but we

78   must do our best to do our homework because it's the homework.

79   Sometimes homework is easy or sometimes it's difficult. If we just do the

80   easy homework, we study, we will not very well.

81   Jenny: Oh it just depends on the topic, I think. If you are interested in one topic,

82   you will do much better. But it is [pause] So why not we change why don't

83   we change role-play and play again.

84   Sara: Play again?

85   Jenny: 'Cause we have a lot of time. [Pause] Okay, which role you prefer to play?

86   Sara: Um, okay, I'm a teacher, you are student now this time uh the teacher tells

87   you you can't change the topic.

88   Jenny: Okay.

89   Sara: Okay?

90   Jenny: Okay, so I will per- persuade you to let me change my topic.

91   Sara: Okay.

              [Students begin a second iteration of the role-play here.]

92   Jenny: Uh, excuse me. Are you professor [pause] Adams?

93   Sara: Yes.

94   Jenny: Uh, my name is Jenny.

95   Sara: Hi Jenny.

96   Jenny: Uh, nice to meet you. Uh how how are you?

97   Sara: Nice to meet you to. [pause] You can have a seat.

98   Jenny: Uh, okay, sure. Uh I just uh I was wondering I am really interested in

99   another topic. Can I change my topic?

100 Sara: Um, I think you can't do that because uh the other student have the same

101   topic. I think if just you change the topic, it's not very well.

102 Jenny: But I I found gradual- I tried to to work on this topic but but gradually I

103   worked the end I found oh it's not suitable it's really not suitable for me to

104   study it. Uh I think if other maybe the topic is sui- sui- suitable for other

105   students, but it's really not suit for me.

106 Sara: But I think uh a little part of student also think this topic is very difficult,

107   but they can uh receive this problems I think you also can. I think you can

108   call others and help you. If you have any problems about the topic, you can

109   ask me, I will help you.

110 Jenny: No no I don't think the topic is very difficult. I just I'm just not interested in

111   it. I'm just I prefer another topic.

112 Sara: But I think uh the homework is not what you are interested in, it's required.

113   So this is your homework, so you must do it.

114 Jenny: No just it just so many topics we can choose, so at ve- at the very

115   beginning I choose this one and I found I am not interested in it later so I

116   want to change another. I think it is available. I think there is so many

117   available topics for me to choose. [pause] What do you think.

[Students talk in L1.]

118   Sara: Okay, you can change it.

119   Jenny: Okay, thank you.

120   Sara: You are welcome.

I will now examine the discourse patterns in the second dyad via the transcript in Example 22. Patterns of interaction in this pair are different from the patterns in the first dyad described above. In Example 22 below, the participants rehearse the role-play most of the time and engage in discussion only for a short period of time. During discussions, they provided brief opinions / feelings about their performance (e.g., lines 88-94: "It sound you will hit me," lines 152-154: "Um yeah it's I should really enjoy. Yeah, I'm enjoy this thing.") and negotiated who will play which role next (e.g., lines 94-97: "So now I'm I'm the professor you are student, right?" lines 195-205: "No you you professor").

Group dynamics between dyad 1 and dyad 2 also differ. In dyad 1, the participants disagreed with each other. In dyad 2, however, the participants were cooperative rather than confrontational. Additionally, in dyad 2 the participant with a pseudonym Sam often stayed off task. For example, in line 45 Sam, in the instructor role, says "Why you come in? [pause] I'm busy," which is clearly inappropriate for an instructor in the U.S. to say. Felix seems to comment about this later during the discussion: "Be (strict?) but polite" (line 90) and "It sound you will hit me. [laugh]" (line 92). Later, in line 152 during the discussion, Sam laughs and says "It's fun," possibly because he is enjoying staying off task and being creative, though not always pragmatically appropriate. In line 209, Sam, acting as a student, again stays off task by being too informal with an instructor "What's up? [laugh]" When Felix, as an instructor, asks Sam how he is doing on his paper without Sam first mentioning the paper (lines 218-219), Sam again stays

137

off task. For example, in lines 220-222, he says "How did you know? … I'm gonna be talking about my topic?" and then he keeps asking about this in lines 228-229. Probably prompted by the fact that Sam keeps staying off task, Felix says "Yeah 'cause I'm professor. I'm the man!" (line 237).

It seems that participants' creativity in dyad 2 prevented them from being able to pay attention to form. For example, even though in the description of the situation the phrase is *the paper is due soon*, Felix keeps using the phrase *due time* (lines 9, 11, 14, etc.). However, after Sam says *due date* (lines 57 and 73), Felix begins using this phrase as well (lines 79, 115, 137, etc.). In line 52, Felix says "Uh I was wondering um if I can change my topic for the paper." The expression *I was wondering* was introduced in class during consciousness-raising. However, in the second rehearsal, Felix makes a request starting with "Um I'm afraid to say this but I need to change my topic" (line 171). That is, Felix might not be able to consistently incorporate the phrases he recently learned due to the fact that he and/or his dyad partner are focusing on being creative.

Just like in dyad 1, participants in dyad 2 do not speak like native-speaking instructors, which is not surprising. For example, Felix says "No sorry, you can't do it because the due time for the paper is soon" (line 14). That is, Felix is being too direct as an instructor, although he is trying to be polite by saying *sorry*. Nevertheless, when Felix is not staying off task, many of his responses as an instructor are not too far off from those of the native speakers. For example, in lines 29-31 his language is indirect (and thus more polite) due to hedges *I don't know*, *maybe,* and *I think*: "Um I don't know but the due time is soon so if you can bring your topic, maybe give you some ideas about this topic. But if you want to change, I think that's too late." Sam, however, is more direct as an instructor: "And then you submit the other topic your topic your

138

your topic you submit to me, and I will give you extra credit" (lines 75-76). In the student role, Felix also uses appropriate arguments for changing the topic of the paper. For example, he says "I tried many time, I can't do anything for this topic" (lines 52-53), "Yeah because I think the original topic I can't get some ideas. Uh I've tried to write more and more, but every time it's be hard to me. [pause] So I know the due date is soon, after two days, but I can do, really, with the new topic" (lines 174-176). Sam's arguments, however, are not appropriate: "This topic is not good not good" (line 21), "The original topic it's kind of bored" (line 122), "Yeah and then not (matched?) about this class" (line 124). While it's possible that Sam uses inappropriate arguments for changing the topic of the paper due to insufficient pragmatic knowledge, it is more likely (based on his comments throughout the dialog) that he is intentionally staying off task. Thus, it is possible that if the participants in dyad 2 stayed on task all the time, Felix would have provided generally good models as an instructor and as a student, and Sam would have been able to pay attention to them and incorporate at least some of them in his production.

Example 22. Transcript of interaction between participants 11 and 151 when rehearsing task 2

1   Sam: Hi Felix.

2   Felix: Hi, how are you doing?

3   Sam: I'm good. Um uh my name is Sam, and I'm your student. Uh do you

4   remember me?

5   Felix: Yeah, you are sitting the back of class all the time, yeah.

6   Sam: Yeah, thank you. But I have a suggestion about your paper.

7   Felix: Yeah.

8   Sam: Yeah. Uh

9   Felix: The due time is soon.

10   Sam: Yeah, I know.

11   Felix: Are you that what you asking about? The due time?

12   Sam: Oh, I see. I just realize now that, but I wanna be a change the topic about my

13   paper.

14   Felix: No sorry, you can't do it because the due time for the paper is soon.

15   Sam: Yeah but in my my opinion uh change the topic is better 'cause I'm preparing

16   the topic that is better than this one. So maybe if I can change the topic I will

17   make a [pause] better paper.

18   Felix: So what's the problem with that with your paper now?

19   Sam: Yeah this this

20   Felix: What is your problem?

21   Sam: This topic is not good not good.

22   Felix: For you or it's not good

23   Sam: No our class.

24   Felix: So you can maybe you can do something for this.

25   Sam: Yeah [pause] I think

26   Felix: Maybe if I can help you a little bit for this topic, not the new one. It's be okay

27   or not?

28   Sam: Yeah.

29   Felix: Okay, we will see. Um I don't know but the due time is soon so if you can

30   bring your topic, maybe give you some ideas about this topic. But if you

31   want to change, I think that's too late.

32   Sam: Oh, too late? So you mean bring

33   Felix: Just bring yeah yeah just bring your paper tomorrow, and I will help you to

34   get some ideas.

35   Sam: Okay.

36   Felix: Okay? Thank you. [long pause] So

          [Students switch roles here.]

37   Sam: Come in.

38   Felix: Uh hi Sam.

39   Sam: Hi.

40   Felix: Hi. How are you doing?

41   Sam: Um how are you?

42   Felix: It's great. How about you?

43   Sam: I'm good.

44   Felix: Uh

45   Sam: Why you come in? [pause] I'm busy.

46   Felix: You are busy now?

47   Sam: Yeah.

48   Felix: I just need one minute.

49   Sam: Just a just a yeah.

50   Felix: Yeah just one yeah.

51   Sam: Okay.

52   Felix: Uh I was wondering um if I can change my topic for the paper because I

53   tried many time, I can't do anything for this topic.

54   Sam: Uhu

55   Felix: It's I don't know it's too close and I don't have many ideas on this topic. So if

56   I can change the topic, and I can do it [pause] quickly.

57   Sam: Uh do you know the due date this uh

58   Felix: Yeah.

59   Sam: day after tomorrow. [pause] Yeah only two days remaining.

60   Felix: Yeah I know.

61   Sam: You know, you know that?

62   Felix: Yeah.

63   Sam: Oh

64   Felix: And if I change uh if I change the topic I can do it.

65   Sam: Uh why you wanna to be uh change the topic?

66   Felix: Uh I was tell you but it's too close for me, and I can't I don't have many ideas

67   for the topic.

68   Sam: Um but the other students have a paper that same topic.

69   Felix: Do you know my topic?

70   Sam: Yeah, I know that.

71   Felix: Okay. I don't know, maybe they have some ideas, but for me, I think it's

72   impossible to get some ideas.

73   Sam: Okay. At first you submit your original topic to me at the due date.

74   Felix: The due?

75　Sam:　Yeah. And then you submit the other topic your topic your your topic you

76　　　　submit to me, and I will give you extra credit.

77　Felix:　So you mean I will give you the original one?

78　Sam:　Yeah.

79　Felix:　And after the due date I will do another one?

80　Sam:　Yeah.

81　Felix:　And if it's good you give me an extra credit?

82　Sam:　Yeah.

83　Felix:　Sounds good.

84　Sam:　How is it?

85　Felix:　Yeah, sounds good.

86　Sam:　Yeah? Okay, thank you.

87　Felix:　Thank you. [long pause]

　　　　　[Students begin a discussion here.]

88　Felix:　Be strict (…?).

89　Sam:　Huh?

90　Felix:　Be (strict?) but polite.

91　Sam:　Yeah.

92　Felix:　It sound you will hit me. [laugh]

93　Sam:　Oh really? [laugh]

94　Felix:　Yeah [laugh] yeah, I don't know. [pause] So now I'm I'm the professor you

95　　　　are student, right?

96   Sam: Yeah.

97   Felix: Okay.

      [Students begin a second iteration of the role play here.]

98   Sam: [pause] Hi professor.

99   Felix: Hi.

100   Sam: Yeah. Hi. Um are you Mister Felix?

101   Felix: Yeah.

102   Sam: I'm Sam. I'm your student.

103   Felix: Nice to meet you, Sam.

104   Sam: Yeah, how are you?

105   Felix: It's great.

106   Sam: Uh I'll say uh do you have a time?

107   Felix: Yeah, yeah. Actually I have meeting after fifty minutes, so it's okay now.

108   Sam: Oh, okay. Uh um I have some suggestion about your topic.

109   Felix: About my topic?

110   Sam: Yeah.

111   Felix: Okay, what do you want to say?

112   Sam: Actually I wanna be uh I wanna change my topic.

113   Felix: Oh

114   Sam: Yeah.

115   Felix: I think that's very hard because you know the due date?

116   Sam: Uhu

117 Felix: It's after two day.

118 Sam: Uhu

119 Felix: So I don't think so if you have uh a long time to change your topic or to write

120   something in your new topic. Uh what's your problem with the original

121   topic?

122 Sam: The original topic it's kind of bored.

123 Felix: For you?

124 Sam: Yeah and then not (matched?) about this class.

125 Felix: Um it's bored, but you know all the learning stuff is bored.

126 Sam: [laugh]

127 Felix: So I don't know. I can help you with your topic if you need some ideas.

128 Sam: Oh sorry I finished this topic my topic yeah.

129 Felix: Oh. So why do you want to change it if you already finish?

130 Sam: Yeah my in MY topic. New topic, new one.

131 Felix: Do you want a new one?

132 Sam: Yeah, I finished my new one.

133 Felix: Oh you finished your new one?

134 Sam: Yeah, yeah, yeah.

135 Felix: Oh, okay.

136 Sam: So (…?) to submit to you.

137 Felix: Okay the due date [pause] okay can you give me tomorrow?

138 Sam: Yeah.

139 Felix: And I can see it if I can accept this or not because

140 Sam: [laugh]

141 Felix: because there is two student came and ask me about the same thing, they

142   want to change the topic but. Okay, you can come tomorrow. And we will

143   talk about this. And if I accept this, don't talk don't tell other your classmate,

144   okay?

145 Sam: Yeah. [laugh]

146 Felix: Keep it secret, okay?

147 Sam: Yeah.

148 Felix: Okay. So good luck, Sam.

149 Sam: Yeah, good luck. [laugh]

150 Felix: And don't forget about tomorrow.

151 Sam: Yeah.

       [Students begin their discussion here.]

152 Sam: [laugh] It's fun.

153 Felix: Um yeah it's I should really enjoy. Yeah, I'm enjoy this thing.

154 Sam: [laugh]

155 Felix: Okay. Now I am a student.

156 Sam: Yeah.

157 Felix: Shall we start?

158 Sam: Yeah.

159 Felix: Okay.

       [Students switch roles here and act out the role-play.]

160 Sam: Come in.

161 Felix: Uh professor, Sam?

162 Sam: Yeah.

163 Felix: How are you doing?

164 Sam: Hi Felix.

165 Felix: Oh you know me?

166 Sam: Yes, I remember you 'cause you are 'cause you you know my class.

167 Felix: Oh, thank you. I'm really appreciate that.

168 Sam: Yeah, how are you?

169 Felix: Uh yeah I'm pretty good. How about you?

170 Sam: Yeah, I'm good.

171 Felix: Um I'm afraid to say this but I need to change my topic, and I'm already done

172   something about the new topic.

173 Sam: Oh, really?

174 Felix: Yeah because I think the original topic I can't get some ideas. Uh I've tried to

175   write more and more, but every time it's be hard to me. [pause] So I know the

176   due date is soon, after two days, but I can do, really, with the new topic.

177 Sam: Um what is your topic?

178 Felix: [pause] Uh I can show it to you tomorrow. Uh you can see. I'm already done

179   with the half of the topic. Maybe if you give me the permission, I can

180   complete it today and give it to you tomorrow to see it, and the day after

181   tomorrow submit it. So what do you think?

182 Sam: Yeah but uh I uh I can't answer in the yesterday 'cause I don't know your

183   topic. When I was check your topic, and you said it's approved or not, then I

184   s- uh I'll tell you about your topic.

185 Felix: So you have a free time tomorrow to come to you?

186 Sam: Yeah, yeah.

187 Felix: Okay. [pause] So it's okay for now to give it to you tomorrow?

188 Sam: Yeah.

189 Felix: Um thank you. I really appreciate this.

190 Sam: Yeah?

191 Felix: Yeah, I'm really thanks.

192 Sam: See you.

193 Felix: So, have a good day.

194 Sam: Yeah.

        [pause] [Students begin their discussion here.]

195 Felix: So it's now ten minutes.

196 Sam: Uhu

197 Felix: So, I'm the professor. Okay, you be the professor, me the student, okay?

198 Sam: No you you professor.

199 Felix: Really?

200 Sam: Yeah.

201 Felix: Oh. 'Cause I'm student. Last time I'm the student.

202 Sam: Yeah, so are stu- you are now professor.

203 Felix: Now I am professor?

204 Sam: Yeah, yeah.

205 Felix: Yeah. That's what I'm talking about. Okay, me the professor. You.

      [Students begin the role play here.]

206 Felix: Come in.

207 Sam: Hi.

208 Felix: Hi.

209 Sam: What's up? [laugh] (This is bad?)

210 Felix: Yeah, it's fine. How are you doing, Sam?

211 Sam: How are you, Felix uh professor?

212 Felix: Uh it's great.

213 Sam: Professor Felix?

214 Felix: Yeah. Oh it's okay to call me Felix only.

215 Sam: Oh really?

216 Felix: Yeah.

217 Sam: Um

218 Felix: So what do you do for now for your paper? Because the due date is after two

219   days, you know?

220 Sam: How did you know?

221 Felix: Huh?

222 Sam: I'm gonna be talking about my topic?

223 Felix: I don't know but

224 Sam: (…?)

225 Felix: No, the due date is soon after two days.

226 Sam: No, no, no [laugh]

227 Felix: You must hurry with this topic. Do you have a problem with your topic?

228 Sam: [laugh] How did you know? How do you know about I will talking about

229   that topic? I never said before you.

230 Felix: Because all the student come today to class and to my office to ask me about

231   the topic.

232 Sam: Oh, really? [laugh]

233 Felix: Yeah.

234 Sam: Maybe that's true 'cause your topic is so hard.

235 Felix: So I'm professor, and I'm always know about what the student think.

236 Sam: Oh yeah. Yeah 'cause professor.

237 Felix: Yeah 'cause I'm professor. I'm the man!

238 Sam: Yeah but you know your topic is too hard to our other

239 Felix: The topic?

240 Sam: Yeah.

241 Felix: Why it's hard? I think all the student doing good with it.

242 Sam: Yeah but in a class with you we understand everything we can understand

243   everything during the class.

244 Felix: You can't?

245 Sam: We CAN can.

246 Felix: Can. Okay.

247 Sam: We can under- easily to understand your class but

248 Felix: The topic you think it's hard to you?

249 Sam: Yeah after the class and then the couple of paper is too hard for us. We

250   cannot

251 Felix: I don't know but I don't know but you must solve this because the due date is

252   soon. And what do you think you can do? You are the student. You want to

253   write in this topic or you [pause] or do you want help with some ideas?

254   'Cause you must hurry. It's soon.

255 Sam: Yeah, I know. Yeah. I want to to write about this topic.

256 Felix: Uhu

257 Sam: I have some the other idea.

258 Felix: What is it?

259 Sam: About my topic.

260 Felix: Yeah.

261 Sam: And I'm already done done my topic.

262 Felix: Done with your topic?

263 Sam: Yeah.

264 Felix: Yeah, that's good.

265 Sam: Yeah that's yeah. So I just ask you about the (prefer?) the submit this topic. If

266   I can or not?

267 Felix: Submit the which topic which topic?

268 Sam: New one.

269 Felix: Oh, you mean new one? I think it's the original topic.

270 Sam: Yeah.

271 Felix: I think it's original topic. But the new one?

272 Sam: Yeah.

273 Felix: I don't know. I think no.

274 Sam: No?!

275 Felix: Yeah, you can't.

276 Sam: Oh

277 Felix: Sorry.

278 Sam: Oh yeah that's

279 Felix: Yeah you must work the original one because I didn't give me the permission

280   to have a a new topic. Because all the student work the same topic.

281 Sam: Oh, that's bad.

282 Felix: So. Do you wanna some ideas to finish your topic because it's really really

283   soon.

284 Sam: Okay, I will try.

285 Felix: Did you done with anythings in the original topic or not?

286 Sam: Yeah yeah, anything.

287 Felix: Nothing or?

288 Sam: Nothing.

289 Felix: Oh

290 Sam: I just tried best on my topic.

291 Felix: I can help it some ideas but I really really want you to finish this soon.

292   Sam: Oh, thank you anyway. Thank you.

293   Felix: Oh, you are welcome.

In summary, CASTs appear to be most beneficial for rehearsal because learners can focus their attention on form and try to make their production more target-like as they rehearse the exact same kind of situation several times. Additionally, native speaker input provided through CASTs serves as models for learners, which they clearly use. Finally, learners appear to stay on task in CASTs. One reason for this could be that native speaker recordings appear naturalistic and prompt learners to behave in the same way they would in real life.

On the other hand, role-plays in learner dyads, while useful, might be less suitable for rehearsal. The utility of role-plays for rehearsal might depend on the partner or group dynamics. If learners in a dyad are not focusing on creativity and make an effort to rehearse the same versions of situations several times, they may also improve on form during pair-work. However, this can be difficult to achieve, as the examples above show. Learners might be staying off task for a variety of reasons: because they want to be creative, because they disagree with the partner, because they are dealing with high-imposition speech acts, which are face-threatening situations, etc. When learners stay off task, their interaction seems to be less pragmatically appropriate. Role-plays are beneficial because through them learners can use language communicatively: learners can negotiate meaning with each other, ask for clarifications, or tap into the creative use of language. However, the analysis above suggests that structured tasks, such as CASTs, may need to be used before role-plays when the goal is to develop pragmatic appropriateness.

5.4 Research Question 3

What are learners' attitudes towards CASTs and pair-work activities?

The results of the Exit Questionnaire (Questions 1 and 9) indicate that 73% of all respondents liked CASTs, 69% liked role-plays, and 57% liked both. A higher percentage of positive responses is desirable if such activities are to be used on a regular basis, and students' comments shed light on how these activities can be improved. Many participants mentioned that while these activities were interesting, they would sometimes lead to boredom as a result of too many rehearsals during the required rehearsal period lasting fifteen minutes. In a non-experimental classroom environment, changing this requirement to a shorter rehearsal period or letting students take control of their learning by rehearsing as much as they feel is necessary to achieve a certain goal, such as a particular grade on an assignment, might result in more positive attitudes.

Based on the data from the Exit Questionnaire (Questions 2, 3, 10, and 11), the majority of students (81 to 85 percent) agreed that both types of activities helped them improve pragmatics and speaking skills, and there did not seem to be a difference in the students' perceptions about the usefulness of these activities for improving pragmatics and speaking skills. Participants understood the value of learning pragmatics, as is evident from their comments: "I face many situation like this almost daily, and sometimes I am not sure how to say this correctly;" "It gives me skills to do things that maybe I didn't do before, also it's help because I'll need it in American culture;" "It really help me to learn how to apologize, and I can use them in the future to talk to American people." Participants also commented on the value of speaking practice: "I'm talking more than I'm used to it;" "I have too many opportunities to say English with others, which improve my speaking skill;" "I believe talking as much as possible is the best way to study language."

Although the majority of learners found both CASTs and role-plays beneficial, they stated a preference for one type of activity or the other when asked to do so (see Exit Questionnaire, Question 15). The most common reason for favoring CASTs was the availability of native speaker input, the value of which for SLA has been firmly established (e.g., N. Ellis, 2009; Gass, 1997; Gass & Mackey, 2006; Krashen, 1982, 1985). In addition to input, exchanges with NSs are also necessary, particularly for the development of pragmatics (Belz & Kinginger, 2002, 2003; Porter, 1986; Sykes, Oskoz, & Thorne, 2008; Thorne, 2003; see Chun, 2008, for a comprehensive review). However, as Lee (2004) observed, low-level learners can feel intimidated when interacting with native speakers. On the other hand, Kuriscak and Luke (2009) found that students who interacted with NSs in a virtual online environment had much more positive attitudes to this type of activity than students who engaged in the same activity but with other language learners. Learners felt empowered that they could successfully communicate with NSs. Kuriscak and Luke believe that in virtual words, learners can communicate with NSs in situations that are less face threatening than in the real word because in the former they can remain anonymous. In this dissertation study, learners engaged in a simulated interaction with pre-recorded prompts rather than in a real interaction with native speakers; however, for some students, these simulations felt close to real. For example, one student commented "This computer program it's like a real conversation," and another person said "This kind of activities are really close to true life." However, other students mentioned that even though they liked CASTs, they also felt the need to talk to actual native speakers. One person wrote "It [CASTs] really help lots. But I think we should talk more with American, would help to learn languages efficiently." Another student's opinion about CASTs was "It's kind of good way to improve our English speaking. But it's better to speak with a native speaker."

Another reason for preferring CASTs was an opportunity for independent self-paced learning. Two participants also commented on the benefit of CASTs as a low-risk environment. One reason learners might feel less apprehensive of their rehearsal via CASTs is that they are not talking to actual people and thus are not being evaluated or judged by anybody, at least at the time of their performance. In general, students thought that CASTs were useful for practice before performing those speech acts in real life. Some comments were "It can help students to do some practice before they meet these kind of normal questions" and "Now I'm not afraid to talk to others." Self-paced individualized learning in a low-risk environment is consistently mentioned as one of the benefits of technology-enhanced instruction (Abras & Sunshine, 2008; Chan & Kim, 2004; Garrett, 1991; Levy, 2009; Sykes, 2008).

Regarding reasons for preferring role-plays, the opportunity to talk to actual people was mentioned most frequently. Students thought that this is more fun or feels more real than responding to pre-recorded prompts on the computer. Interestingly, many of those who preferred CASTs felt that CASTs were close to real life. These comments seem to reflect personality differences such that students who like to work independently tend to enjoy CASTs, while individuals that like to work with others prefer real face-to-face communication. That is, students' attitudes might depend on whether they are introverts or extraverts. "The stereotype of an introvert is someone who is much happier with a book than with other people. The stereotype of the extrovert is the opposite: someone happier with people than with a book" (Gass & Selinker, 2008, p. 433). Extraversion/introversion dimension is included in several models of personality traits (e.g., Eysenck & Eysenck, 1985; McCrae & Costa, 2003) and has been one of the most-researched traits in SLA (Dörnyei, 2005). Introverts might study better, and extroverts engage in more talking and social activity, both of which are necessary for language learning (Gass &

156

Selinker). Thus, instruction should be provided in different ways in order to accommodate various individual differences.

Although an introversion/extraversion dimension of personality appears to play a large role in students' attitudes towards CASTs and role-plays in pairs, there might be other mediating factors. For example, several students who initially stated that they prefer to study with somebody rather than alone indicated in the Exit Questionnaire that they prefer CASTs.[30] One of such individuals wrote on the background questionnaire "If in the something, you came across in problem, we can help each other. We are only speak English, it can help me improve my English. If by myself, no one to help you and nobody to talk with you." On the Exit Questionnaire, this student wrote "By myself cannot nervous." At the beginning of the semester, this individual might have not considered that speaking practice could take place via CASTs because he/she was not aware of them. Additionally, students' attitudes towards CASTs might be influenced not only by personality traits, but also by the task: when performing face-threatening speech acts like the ones used in this study, many students who prefer to study with others might nevertheless favor the low-risk environment of CASTs. Other students who initially preferred to work with others kept their position and favored role-plays at the end of the course, perhaps because they were not convinced that CASTs are close to real life. More detailed survey instruments and interviews are necessary to investigate the exact underlying reasons for students' attitudes.

Of those who preferred pair-work activities, several individuals mentioned the benefits of interaction. These included opportunities to ask for repetition or clarification, to point out mistakes to each other, and to learn new phrases from their pair-work partners, all of which have been recognized as the benefits of interaction in prior research (see Gass, 1997; Gass & Mackey,

---

[30] Such comparison could be done only for about half of respondents because several students chose to provide anonymous responses on the Exit Questionnaire.

2006; Long, 1996; Mackey & Goo, 2007). In the present study, some students also liked the fact that they could talk to interlocutors at the same level of proficiency. For example, one participant said "It helps me to interact with my classmate and my classmate speaks English like me slowly."

Finally, it also seems that participants' preferences could be in part influenced by their group assignment: students tended to prefer the type of activity to which they had more exposure. If students are equally exposed to both types of activities, fewer learners might have a definite preference for one activity or the other. This issue is worth investigating in the future because it provides implications for classroom use.

Participants' reasons for their preferences align with empirical research on the benefits of technology-enhanced instruction, authentic input, and interaction. It is also encouraging that even though students had preferences for one type of activity over another, more than half of the students enjoyed both activities, while the rest of the students did not dislike these activities but rather remained neutral to them. This finding informs the implications for language learning discussed below. Students' comments, such as those pertaining to the length of rehearsal period, can also inform future design and use of CASTs for language learning.

CHAPTER 6: CONCLUSION

6.1 Summary of Findings

The findings of this study can be summarized as follows. The results of mixed-design ANOVAs indicate that all groups improved on both rehearsed and new tasks only in terms of vocabulary diversity. All groups also improved on pragmatics on new but not on rehearsed tasks. First, these findings indicate that when performing complex communicative tasks, such as high imposition speech acts, learners focus on improving informational content rather than fluency, accuracy, and syntactic complexity of oral responses. Second, pragmatic development may not be apparent immediately even when pragmatic appropriateness is the focus of the tasks.

In this study, the prediction that the structured task group would outperform the pair-work group was not borne out for any of the measures, although there was a trend such that only the structured task group improved on pragmatics on the rehearsed task. If there are differences between rehearsal via CASTs versus in learner-learner dyads, the use of a more interactive computer program that allows for the provision of feedback might be necessary to uncover them.

While the experimental groups had the same learning outcomes, their discourse patterns during task rehearsal differed. These discourse patterns suggest that CASTs allow learners to focus on form due to several similar rehearsals, and learners are able to incorporate input from CAST prompts into their subsequent production. However, discourse patterns in dyads differ depending on the dynamics of the pair. Learners may become more pragmatically inappropriate when they stay off task. They may stay of task for a variety of reasons, such as because they disagree with each other or because they want to be creative. If learners are creative during the role-play rather than attempting to rehearse the same version of the task, they focus less on form.

The results do not provide conclusive evidence that task rehearsal is beneficial because the control group did not differ from the experimental groups on gains. However, the relative shortness of pragmatic instruction, lack of immediate feedback, access to pragmatics in the second language environment, and low proficiency level of the learners may have lessened the effect of task rehearsal on pragmatic development.

The questionnaire data indicate that while 57% of respondents enjoyed both CASTs and pair work, most students favored one type of activity or the other partly due to their personal preferences. The findings are discussed in terms of implications for using CASTs in face-to-face, hybrid, or online courses.

There are also important findings that inform the design of CASTs. First, especially for the learning of pragmatics, extensive consciousness-raising phase should precede rehearsal via CASTs. Unless timely feedback is provided, students are likely to use unfamiliar input from CASTs incorrectly because they may not be able to decipher pragmatic meanings. The questionnaire data indicate that students who prefer to study alone tend to enjoy CASTs more than role-plays in pairs, at least when the tasks are face-threatening speech acts. On the other hand, those who like studying with others tend to prefer role-plays in pairs. However, there is also some indication that students tend to prefer those activities to which they were exposed the most. Thus, it is possible that many students would not have a preference if they were equally exposed to both types of activities. Generally positive attitudes to both types of activities and students' agreement that both activities improve pragmatics and speaking skills suggest that CASTs and role-plays in pairs can be successfully utilized for practicing these skills.

6.2 Limitations and Future Directions

Some limitations of this study need to be acknowledged. These limitations concern (a) measures used, (b) the design of CASTs, and (c) the design of the study.

One limitation of this study is that the measures used to assess changes in pragmatics and oral development might not have been sensitive enough for finding group differences, and it is thus advised that future studies investigate a variety of measures before definite conclusions are drawn on the effect of CASTs. For pragmatics, the rating scale ranged from 1 to 6. Although both raters found this scale optimal, the range might nevertheless be too narrow for uncovering fine-grained changes. Magnitude estimation method, first introduced to SLA by Bard et al. (1996) for use with acceptability judgments, might be another way to look at small changes. Alternatively, assigning separate scores for the amount of information provided, formality, directness, and politeness (as in Hudson et al., 1995) and then combining the scores on these four aspects rather than assigning one overall score for pragmatic appropriateness might also help unveil smaller changes in pragmatics (K. Walters, personal communication, February 3, 2011). On the other hand, examination of pragmatic development through a discourse analysis approach rather than through ratings might be more fruitful when small changes are evaluated (K. Bardovi-Harlig, personal communication, March 27, 2011). Kasper (2004, 2006) argues that discursive approach is most applicable to the analysis of speech act pragmatics. As the raters in this study have noted, one inappropriate statement from a participant may significantly influence the rating even if the rest of the response is appropriate, in which case learners would not always receive much credit for the use of polite phrases and strategies. Thus, investigating changes in the use of particular pragmatic strategies, for example, the use of mitigators, downgraders, and hesitations in request speech acts, might be more adequate for examining the effects of CASTs and role-play activities on pragmatic development. Taguchi (2006) is the only study that

compared ratings of pragmatic appropriateness with the participants' use of discourse strategies. Ratings of pragmatic appropriateness were adequate for distinguishing between learners of lower and higher proficiency in her study, but it should be investigated whether the analysis of discourse strategies is more sensitive to small changes in the learners' pragmatic development than appropriateness ratings. Other specific changes in pragmatics to be examined could be routine formulae, or in other words, formulaic expressions (Bardovi-Harlig, 2006, 2009). Such detailed analyses might be useful not only for measuring small changes, but also for investigating whether rehearsal via CASTs and role-plays in pairs helps develop pragmatics in general or only certain aspects of pragmatics.

There are also issues with the objective measures of fluency, complexity, and accuracy. While they are frequently used in SLA research in general, and in planning studies in particular, more research is needed to perfect the assessment of these areas in terms of reliability, validity, and sensitivity of the measures (Housen & Kuiken, 2009). Segalowitz (2010) states that although measures of fluency have been heavily researched, there is still no consensus as to which measures truly reflect learners' ability in this area. While it has been confirmed that pause phenomena, especially within-clause pausing, influence human ratings of fluency, one of the numerous challenges is deciding on the minimum length of pauses when calculating pause-related measures. Iwashita et al. (2008) point out that measures of complexity, such as the T-unit complexity ratio and the dependent clause ratio, are also problematic even though they had been recommended on the basis of previous studies as among the most useful measures of complexity (Wolfe-Quintero et al., 1998). While measures based on subordination are applicable to long stretches of written and oral discourse, they might need to be modified for oral interaction where turns are short. In an effort to make complexity measures originally developed for written data

162

more applicable for oral data, Foster et al. (2000) have proposed an AS-unit, which is similar to the T-unit, but is adapted to oral and highly interactive discourse. Rather than searching for perfect measures, researchers may need to identify the most appropriate measures for a particular case. For example, when examining syntactic complexity, the choice of measures can depend on a variety of factors, such as the kinds of tasks used in research (Norris & Ortega, 2009) and learners' proficiency level (Norris & Ortega; Shea, 2011). Regarding accuracy, both global measures, such as the ratio of error-free clauses or the number of errors per 100 words, and errors on specific aspects have been used for research. When the goal is to find small changes, the examination of grammatical errors of certain types might be most effective. For example, Gass et al. (1999) examined grammatical improvements on Spanish verbs *ser* and *estar*; Mochizuki and Ortega (2008) measured the degree of accuracy of relative clauses; Wigglesworth (1997) examined the suppliance of plural –s and the accuracy of verbal morphology and indefinite articles.

It might also be inappropriate to investigate improvement through the combination of pragmatic ratings and measures of fluency and complexity due to their interaction. In high-imposition speech acts, it is often appropriate to use hesitations as downgraders (e.g., Cohen, 2005; House & Kasper, 1981). In the present study, the raters also commented on the use of hesitations as polite strategies. Thus, as pragmatic ratings increase based on appropriate use of hesitations, fluency decreases. That is, learners might be disfluent deliberately if they are hesitating for pragmatic purposes. Improvement in pragmatics might also misleadingly show improvement in syntactic complexity. For example, if language learners use such newly acquired formulaic expressions as *I was wondering if* and *if you don't mind*, the syntactic complexity of their responses (measured in clauses per AS-unit) would increase (K. Bardovi-Harlig, personal

communication, March 27, 2011), but it does not necessarily mean that overall, the language that learners produce is more syntactically complex. That is, learners might not be able to analyze newly acquired formulaic expressions and extend the use of subordinate clauses to other contexts.

Another limitation of this study is that while feedback is crucial, it could not be provided through the *Conversations* program, which was used to deliver CASTs. However, with rapid technological developments, it might soon be possible to make CASTs adaptive to learners' responses and to provide interactive feedback. While voice recognition technology is not sophisticated enough at this point to be applied to language learners (Derwing et al., 2000), according to Morton and Jack (2005, 2010), speech recognition becomes more accurate when restricted scenarios are used because learners' responses are somewhat predictable in such situations.

For an in-depth understanding of the process that learners go through when rehearsing via CASTs verbal reports can be employed. For example, following stimulated recall methodology (Gass & Mackey, 2000), learners would be instructed to watch the video of their rehearsal via CASTs and provide any comments on what they were thinking of at the time of their performance of the task. This methodology would be particularly valuable for investigating what learners attend to in the input when using CASTs. Interviews can be used to have a better understanding of what learners' attitudes are guided by.

One more area for future directions is investigating the impact of CASTs on speaking and especially on pragmatics in a foreign language environment. It is possible that all groups in this study, including the control group, improved on pragmatics by the end of the semester because they needed to use this skill not only in their classes, but also in their daily life in the U.S., as some of the experimental group participants mentioned in the Exit Questionnaire. Future studies

might reveal that CASTs are most beneficial in a foreign language environment where students have fewer opportunities to practice pragmatics.

Finally, because this study is exploratory in nature and because the goal was to maximize pedagogical utility of CASTs, it was not investigated whether CASTs are beneficial because they provide structured guidance to learners, authentic input, or both. Future research will be needed to disentangle the variables of structured rehearsal and authentic input. However, the examination of discourse patterns seems to suggest that both factors contribute to language learning.

6.3 Pedagogical Implications

This study provides important implications for second and foreign language instruction with the focus on pragmatics. The results support the idea that pragmatics is a difficult aspect of an L2 and takes a long time to develop (Blum-Kulka & Sheffer, 1993) because learners did not improve on pragmatics immediately after rehearsal, but they did improve on pragmatics on new tasks tested at the end of the study. Several examples from the data in this study also show that students' knowledge of pragmatics is in a developmental stage. Thus, while task rehearsal generally improves subsequent oral production by reducing attentional burden (as has been shown with vocabulary diversity in this study), it may not have a large effect on pragmatics, especially for learners at a low proficiency level. For this reason, in pragmatics instruction <u>extensive</u> consciousness-raising needs to take place before production. The data in this study also support the claim that explicit teaching of pragmatics is necessary (Bardovi-Harlig, 2001; Jeon & Kaya, 2006; Kasper & Rose, 2001; Rose, 2005) because while learners may notice new forms in input and use them, they can incorporate them inappropriately in specific contexts. That is, learners first need to be taught (deductively or inductively) in which pragmatic contexts certain

165

forms are applicable. Likewise, feedback is a necessary part of pragmatics instruction at least because when dealing with learners from different linguistic and cultural backgrounds, a teacher will not be able to foresee all pragmatic mistakes that learners will make and draw learners' attention to these aspects during consciousness-raising. Although consciousness-raising and feedback are part of various models of pragmatic instruction (e.g., Cohen, 2005; Félix-Brasdefer, 2006; Judd, 1999; Olshtain & Cohen, 1991), the importance of these elements needs to be stressed.

In this exploratory study of CASTs, quantitative analysis of gains suggests that task rehearsal via CASTs is as effective as rehearsal through role-plays in pairs because the two experimental groups did not differ in gains on any of the measures of pragmatic and oral development. However, the qualitative analysis of discourse patterns suggests that CASTs and role-plays in dyads may offer different benefits. CASTs appear to be more conducive to focusing on form because prompts in CASTs do not change, and thus learners are more likely to rehearse the same version of the task several times. CASTs can also provide models of authentic input, which learners tend to incorporate into their subsequent production. (But again note that learners can incorporate some input incorrectly in terms of pragmatics, and thus feedback is necessary). However, when the goal is for students to practice pragmatics communicatively, role-plays are more appropriate. The advantage of role-playing in pairs is that learners can be creative and negotiate meaning. However, teachers should keep in mind that when working with each other, learners may stay off task, which in some cases leads to pragmatically inappropriate behavior. On the other hand, in CASTs learners tend to stay on task, possibly because they feel closer to taking part in a particular situation instead of <u>pretending</u> to be in that situation. That is, the appearance of native speakers in video-recorded prompts may make this impression on learners.

Thus CASTs may be particularly useful for rehearsing speech acts, which follows Judd's (1999) suggestion that in pragmatics instruction structured practice should take place before open-ended production. It should be also kept in mind that learners display similar discourse patterns in CASTs, but discourse patterns in role-plays in pairs differ. Interactions in dyads can be affected by group dynamics or learners' proficiency level (Cao & Philp, 2006; De Saint Léger & Storch, 2009; S.-J. Kang, 2005; Kowal & Swain, 1997; Watanabe & Swain, 2007).

Assuming that extensive consciousness-raising activities are provided first, CASTs are particularly useful for independent self-paced practice: for example, in online courses when face-to-face interaction might be less feasible or less frequent than in traditional courses or in a foreign language environment when learners generally interact with speakers of the same L1. CASTs could be also assigned as homework to prepare students for in-class interaction. This preparation can free more class time for speaking (Zyzik & Polio, 2008). Learners should first receive feedback on their production in CASTs before they engage in open-ended practice in class.

Self-paced rehearsal via CASTs provides a low-risk environment, and some participants commented that they are less nervous when rehearsing via CASTs than via role-plays in pairs. This could be due to the fact that nobody evaluates learners' performance (at least immediately) in CASTs, while some students may be apprehensive of talking to their partners, especially if they are in non-cooperative dyads. However, as participants' comments on the Exit Questionnaire suggest, personal factors are at play because other students felt quite comfortable talking to their peers. In fact, about half of the students preferred role-plays over CASTs, the reasons being that these students like talking to real people and that they recognize the benefits of face-to-face interaction with their peers.

It is advised that both CASTs and role-plays in pairs are used for learning pragmatics, but CASTs should precede role-plays, until other studies are conducted to investigate CASTs in more depth. The reasons are that (1) there were no significant differences in gains between the two experimental groups; (2) discourse patterns show that CASTs and role-plays in pairs have different benefits, and (3) although most participants expressed a preference for either CASTs or role-plays when they were asked to do so, the students generally enjoyed or at least did not dislike the less-preferred type of activity. Both activities should be used because some learners prefer to engage in self-paced practice, while others prefer talking to real people.

With regard to the effect of rehearsal via CASTs and role-plays in pairs on oral development, it appears that only vocabulary diversity improves. That is, because learners performed communicative tasks, they were probably focusing their attention on developing content than on improving the structure of their message and the fluency of their production.

6.4 Contribution of the Study

To conclude, the main contribution of this study lies in exploring the potential of rehearsal via CASTs for language development, and specifically for the development of pragmatics and oral skills. Given that there is a need for materials and methods that can be used for independent and online learning, this study adds one more piece to the puzzle on the possibilities afforded by modern technologies. To my knowledge, the use of CASTs for language learning has not been investigated before. Thus, while some questions remain unanswered, the contribution of this study is in providing future directions in terms of the design of CASTs and methodology for evaluating their effectiveness.

Another contribution of this dissertation lies in examining the effect of task rehearsal on pragmatic development since this issue has not been investigated before. The main finding is that

pragmatic development might be measurable only after some period of time rather than immediately after treatment. While this finding is not conclusive, it provides suggestions for future research in terms of examining pragmatic development via a variety of measures and methodologies.

APPENDICES

APPENDICES

Appendix A: Background Questionnaire

1. Name: _____
2. Age: _____
3. Gender: ___ Male   ___ Female
4. First/native language(s): _____
5. Where have you studied English and for how long?
☐ Elementary School   Number of years/months: _____
☐ Middle/High School   Number of years/months: _____
☐ Private tutoring   Number of years/months: _____
☐ Language School   Number of years/months: _____
☐ College/University   Number of years/months: _____
☐ ELC at MSU   Number of semesters: _____
☐ Other _____   Number of years/months: _____

6. Which English-speaking countries have you been to?

Country	For how long	For what purpose

7. How comfortable are you with using technology (computers, Internet, etc.)?
☐ very comfortable
☐ comfortable
☐ somewhat comfortable
☐ not comfortable
☐ very uncomfortable

8. What kinds of technology have you used for language learning?

Technology	Did you like it? Why?

9. Do you prefer learning a language with technology or without? Why?

10. Do you prefer studying by yourself or with someone? Why?

11. Do group activities in class help you learn a foreign language? Why?

Appendix B: Situation Descriptions for Rehearsal and Pair Tests and CAST Prompts for Rehearsal

Request Task 1

Role X
Imagine that you and your friend are hanging out in a coffee shop right now. Ask him to help you move this weekend. You really need his help. *Be polite.*

CAST Prompts[31]
1. [Participants see a person typing on a computer]
2. Well, what do you have to do? I mean like, how long is it gonna take? Because my sister is coming to visit this weekend.
3. Um, I guess I could stop by Saturday before she gets here or maybe Sunday after she leaves.
4. Okay, I'll call you Friday and tell you what time I can come by.

Role Y
You are with your friend now. Your friend is asking you to help her move this weekend, but you already have other plans. You will try your best to help your friend, but it is difficult for you. *Be polite.*

CAST Prompts
1. Hey, what are you doing this weekend? Wanna help me move?
2. I'm moving both Saturday and Sunday. If you can come either of those days anytime at all, I'd appreciate any help.
3. I mean if you can't come, I understand, that's okay.
4. Thanks. I'd really appreciate it if you would be able to help.

---

[31] Participants only heard CAST prompts; they did not see them in a written form. CAST prompts were presented in the order they are listed.

Request Task 2

Role X
You are a student. Ask your instructor if you can change the topic of your paper. The paper is due soon. You <u>realize</u> (= you know) that you do not have a lot of time to work on the new topic. However, you have good reasons for changing the topic of your paper. Try to get permission from your instructor to change the topic. *Be polite.*

CAST Prompts
1. How can I help you?
2. Wow, it's pretty late to be changing your topic. If if you'd come to me a few weeks ago, maybe we could have worked something out. But uh I really do think you should stay with your original topic.
3. Yeah, I understand. But my chief concern is that we've already met about the previous topic and I've looked at outlines for the class. So if you did change your topic now, you'd be handing in a paper without any of that previous guidance.
4. Okay, so you can change your topic, but I think you are taking a big risk because I'm not going to have much time to help you with it since we are approaching the end of the semester. In fact, if I were you, I would seriously consider whether you wanna do this or not.
5. Okay, either way, just let me know what you decide.

Role Y
Imagine that you are an instructor. Your student asks you if he or she can change the topic of the paper. But you think that it is too late to change the topic now because the paper is due soon. Try to convince the student to write the paper on the original topic. *Be strict but polite.*

CAST Prompts
1. Hi. I was wondering if I could switch topics for my paper. I know it's due soon. But I'm not finding very much stuff on this topic and there's another topic that I can find a lot of information on.
2. I'm really really having problems with my first topic and I've already found a lot of stuff on the second topic and you know like if I can handle switching topics, would you allow it?
3. I understand but I really really like this new topic. So what if I come up with an outline and get it to you tomorrow?
4. Yeah, I'll think about it some more. And when I make a final decision, I'll let you know. Thanks.

Refusal Task 3

Role X
You have just finished talking to your academic advisor about your courses. Now your advisor <u>strongly suggests</u> you do an internship in the summer, but you are <u>reluctant</u> to do it (= you do not want to do it). This internship is <u>not required</u> (= you do not have to do it). *Reply to your advisor politely.*

CAST Prompts
1. Also, there is one other thing I wanted to talk to you about. It looks like there's a possibility for an internship in Washington DC. I think you'd be a great candidate. Would you be interested?
2. Okay, I understand how you feel. But I do think this is a great opportunity. Uh, you can get some experience, make some connections, and just have fun in Washington.
3. Okay, the application is due in two weeks, so you are going to need to make a decision soon. I'll forward you some more information on this. It is unpaid, so money might be an issue, but we will help you find accommodation and you might ask the internship director for a small compensation.
4. Alright, if you have any questions, just let me know. And I'd be more than happy to write you a letter of recommendation.

Role Y
Imagine that you are an academic advisor. You just finished talking about courses with one of your students. Now you <u>strongly suggest</u> to your student to do an <u>internship</u> (=a paid or unpaid job) this summer in Washington, DC. This internship is <u>not required</u> (=the student doesn't have to do it), but you think that it is very important. Explain to the student why this internship is important and try to convince the student to do this internship. *Be polite.*

CAST Prompts
1. Okay, so looks like I'm right on track.
2. Oh, I really appreciate the fact that you consider me a good candidate for this internship, but I was kind of planning on doing something where I could make some money over the summer, so I'm not too sure.
3. I mean, you are probably right. And I really think it is a good idea that I get an internship in before I graduate, but I kind of need to work and make money this summer so that that way I can continue to study here in the fall. But let me think about it a little bit first.
4. I'll definitely think about it. And yeah, thanks for letting me know.

Appendix C: Situation Descriptions and CAST Prompts for Computer Tests

Request Task 1
Ask your friend if you can stay at his/her parents' place in New York. You have an interview in New York. This is very important for you. *Be polite*.
IMPORTANT: You should ask about this <u>right after</u> your friend says "Hey, what's up?"

CAST Prompts
1. Hey, what's up?
2. Uh, I mean, congratulations. It's great that you got that interview and all. And I'd love to help, but I'd have to ask my parents first because their apartment's kinda small.
3. Yeah, I don't know if they are gonna be home. They kinda go out of town a lot, but I'll call them tonight and let you know.
4. Yeah, sure, no problem. I'll let you know what they say.

Request Task 2
Ask your instructor for help on homework. The instructor is busy and even had to cancel office hours this week. However, this homework is very difficult for you and you cannot do it without help. *Be polite.*

CAST Prompts
1. Okay, what's your question?
2. Oh, I'm I'm really sorry, this week is very bad for me. I'm extremely busy. And in fact that's why I had to cancel my office hours. What I'd recommend is that you talk to the tutor. He is very good at explaining these things.
3. I'm afraid I'm really pressed for time right now and in fact I've gotta go teach in a few minutes. And and I truly apologize but this week as a said is really busy. I'm I'm leaving town, so I'm not gonna be able to meet with you in person. But if you'd like to email me, I'd be more than happy to to respond to you. And also you might wanna check the PowerPoints that we used in class. You could also use the software in the language lab.
4. Yeah, and we can talk during my office hours next week.

Refusal Task 3
You came to talk to your advisor about the classes you should take in the spring semester. Your advisor *strongly suggests* that you take a summer course because a great visiting professor will be teaching it. However, <u>you don't want</u> to take any summer classes. This summer class is <u>not required</u> (= you do not have to take it). *Be polite*.

CAST Prompts
1. Okay, this looks like a good plan, but for the sociolinguistics course I'd recommend that you actually wait until the summer to take it. Because this summer we have a visiting professor from Harvard and she teaches a great course in sociolinguistics, so why don't you take that one.
2. Okay, that's really too bad because this will be your only chance to take class with professor Davis. And I do think the sociolinguistics course from her will look really good on your transcript. Also, the summer course is only sorry the summer session is only seven weeks long, so you know if you wanted to work or travel, you'd still have plenty of time to do that.

3. Okay, I'll forward you the information on the course, and you can decide if this is something you are interested in. But I have to say: I think this is a once-in-a-lifetime opportunity.

Appendix D: Transcripts of Videos Used in the Consciousness-Raising Phase

Request Video 1
Brian: Hey Gary, how's it going?
Gary: Um, okay I guess.
Brian: Um, I was wondering if you could look over my paper for my history class?
Gary: Um, what did you want me to to check?
Brian: Um, just kind of look it over and see if it makes sense and see if it kind of flows right and that kind of thing.
Gary: Uh, and when do you need it by?
Brian: Um, it's due on Friday afternoon.
Gary: Friday afternoon. Uh well, I guess I could try you know I'm actually trying to finish up with packing and moving so.
Brian: Oh, that's right, yeah
Gary: But uh I can definitely try though I can't promise.
Brian: Okay, yeah you don't have to go through you know like every single word but if you can just kind of get your feel for it that'd be great.
Gary: Okay, I'll try my best to do it.
Brian: Okay, great, thanks a lot.

Request Video 2
Student: Uh, excuse me, are you Dr. Johnson?
Professor: Yes, I am.
Student: Um, my name is David.
Professor: Hi.
Student: Hi, how is it going?
Professor: Nice to meet you. Have a seat.
Student: Okay, thanks.
Professor: What can I do for you?
Student: I was wondering, I'm really interested in getting into your um biology class that's being offered uh this spring.
Professor: Yes.
Student: But the class is full.
Professor: Yes, it's completely full, there are actually already 30 people in the class.
Student: Yeah. So, I was wondering if there's any chance you'd maybe if you ever do overrides 'cause I'm really interested in this class, um, I think I'd be a really good student. Also, I should mention that um I need it for my major… and I haven't taken this class yet.
Professor: I see. Um, what year are you?
Student: I'm a senior.
Professor: Oh, so you'll be graduating this year?
Student: Yeah, I'm set to graduate I just need this class.
Professor: I don't know because usually I don't like to add any students that are like over the limit of 30 [yeah] also because the classroom space is kind of limited [yeah] Um, goodness. Well, actually I think let's do it this way for now uh usually some students drop the course so if you can just come to the course for the first week and

we'll see whether or not anyone drops and if they drop then you can kind of take their space in the class.

Student: Yeah, okay.

Professor: Um, if that would work for you.

Student: Yeah.

Professor: I can't promise you right now that I will write you an override but just come for the first week of class because I really think somebody might drop it.

Student: Okay, great it helps I really really I have to get in the class.

Professor: I understand that it's important so we'll try it that way and see whether or not that works then we'll talk later if nobody does drop okay?

Student: Okay, great, thank you so much.

Professor: No problem. I'll see you in class then on Monday.

Student: Yeah.

Refusal Transcript 3 (no video)

Professor: Alison, great work on the final paper!

Student: Oh, thank you very much.

Professor: Well, I was thinking would you be interested in presenting this paper at a conference that our department will organize next semester? It would be an excellent opportunity for you to present a paper at a conference if you are still thinking about going to graduate school. And this would also benefit our department.

Student: I would love to. When exactly is this conference?

Professor: It's the week before finals.

Student: Oh, then I'm not quite sure whether I can do this. My husband and I were kind of thinking about going to Egypt to see his family. Next semester I'm taking an independent study and a couple online courses, so I could leave before the semester ends.

Professor: Sure, I understand. But I do think this is a great opportunity and you definitely ought to discuss this with your husband. I suggest you explain to him how important this opportunity is for you.

Student: Um, well, I suppose we could postpone the trip, but I'd have to talk to him.

Professor: Definitely, first see what your husband says, but the application is due in two weeks, so had better make a decision soon. I can help you with the application if you decide to go for it.

Student: I will let you know in a day or two. And thank you so much for letting me know. I appreciate it.

Appendix E: Exit Questionnaire

I will see your answers only after you receive your final grade in this class.

If you don't want to, you don't have to write your name.

Name (optional): _____

I would like to know what you think about different activities that you did in class.

Your opinions will be very helpful for me and for ELC (but only I will see your answers).

Please *circle your opinion* and *write your explanations or comments*.

Part A: Online activities (talking on the computer)

Apologizing, asking for something, and saying "no"

1. I liked these activities.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

2. These activities helped me learn how to apologize, ask for something, and say "no."

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

3. These activities helped me improve my speaking skills.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

4. I was nervous doing these activities.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

5. I liked using this computer program.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

6. I like using computers for learning languages.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

7. I liked "*Talking to Jenny*" activities.[32]

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

8. What did you learn from online activities?

_____

Part B: Role-playing in pairs

9. I liked doing role-plays in pairs.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

10. Role-plays helped me learn how to apologize, ask for something, and say "no."

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

11. Role-plays helped me improve my speaking skills.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

12. I was nervous doing these role-plays.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

13. I didn't like *recording* role-plays on the computer.

strongly agree   agree   neutral/in the middle   disagree   strongly disagree

Please explain _____

14. What did you learn from role-plays?

_____

15. Which activities did you like *more*?

_____ talking to a computer (online activities)

_____ doing role-plays in pairs

Why? _____

Thank you very much!

---

[32] Students used CASTs to respond to prompts on everyday topics, such as their hobbies and interests. These activities focused on speaking skills in general rather than specifically on pragmatics. The data from these activities were not analyzed as they were part of the course but not part of the research study.

REFERENCES

REFERENCES

Abras, C. N., & Sunshine, P. M. (2008). Implementing distance learning: Theories, tools, continuing teacher education, and the changing distance-learning environment. In S. Goertler & P. Winke (Eds.), *Opening doors through distance language education: Principles, perspectives, and practices* (pp. 175-201). San Marcos, TX: CALICO Publications.

Austin, J. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing, 19*, 453–476.

Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

Bard, E., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language, 72*(32-68).

Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics? In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 13-32). Cambridge: Cambridge University Press.

Bardovi-Harlig, K. (2002). A new starting point? Investigating formulaic use and input. *Studies in Second Language Acquisition, 24*, 189-198.

Bardovi-Harlig, K. (2006). On the role of formulas in the acquisition of L2 pragmatics. In K. Bardovi-Harlig, C. Felix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning 11* (pp. 1-28). Honolulu, HI: University of Hawai'i press.

Bardovi-Harlig, K. (2009). Conventional expressions as a pragmalinguistics resource: Recognition and production of conventional expressions in L2 pragmatics. *Language Learning, 59*(4), 755–795.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition, 11*(1), 17–34.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*, 233-262.

Bardovi-Harlig, K., & Hartford, B. (1993). Learning the rules of academic talk: a longitudinal study of pragmatic change. *Studies in Second Language Acquisition, 15*, 279-304.

Bardovi-Harlig, K., & Hartford, B. (1996). Input in an institutional setting. *Studies in Second Language Acquisition, 18*(2), 171-188.

Belz, J., & Kinginger, C. (2002). The cross-linguistic development of address form use in telecollaborative language learning: Two case studies. *The Canadian Modern Language Review, 59*(2), 189-214.

Belz, J., & Kinginger, C. (2003). Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms. *Language Learning, 53*(4), 591-647.

Bhat, S., Hasegawa-Johnson, M., & Sproat, R. (2010, September). *Automatic fluency assessment by signal-level measurement of spontaneous speech.* Paper presented at the INTERSPEECH Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology, Tokyo, Japan.

Biesenbach-Lucas, S. (2007). Students writing emails to faculty: An examination of e-politeness among native and non-native speakers of English. *Language Learning & Technology, 11*(2), 59-81.

Blum-Kulka, S. (1997). Discourse pragmatics. In T. Van Dijk (Ed.), *Discourse as social interaction* (pp. 38-63). London: SAGE.

Blum-Kulka, S., House, J., & Kasper, G. (Eds.). (1989). *Cross-cultural pragmatics: Requests and apologies*. Norwood, New Jersey: Ablex.

Blum-Kulka, S., & Sheffer, H. (1993). The metapragmatic discourse of American-Israeli families at dinner. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage Pragmatics* (pp. 196-223). Oxford: Oxford University Press.

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing, 26*(3), 341–366.

Brown, A., Iwashita, N., & McNamara, T. (2005). *An examination of rater orientations and test taker performance on English for academic purposes speaking tasks* (TOEFL Monograph Series No. 29). Princeton, NJ: Educational Testing Service.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Bruton, A., & Samuda, V. (1980). Learner and teacher roles in the treatment of oral error in group work. *RELC Journal, 11*, 49-63.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Harlow, UK: Pearson.

Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37-74). Amsterdam: John Benjamins.

Bygate, M., Skehan, P., & Swain, M. (2001). Introduction. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 1-20). Harlow, UK: Pearson.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 2-27). London: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behaviour in whole class, group and dyadic interaction. *System, 34*, 480–493.

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. Alcón Soler & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 41-58). New York: Springer Science + Business Media LLC.

Celce-Murcia, M., Dőrnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics, 6*(2), 5-35.

Chan, W. M., & Kim, D.-H. (2004). Towards greater individualization and process-oriented learning through electronic self-access: Project "e-daf." *Computer Assisted Language Learning, 17*(1), 83– 108.

Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge: Cambridge University Press.

Chiu, T.-L., Liou, H.-C., & Yeh, Y. (2007). A study of web-based oral activities enhanced by Automatic Speech Recognition for EFL college learning. *Computer Assisted Language Learning, 20*(3), 209 - 233.

Chun, D. (2008). Computer-mediated discourse in instructed environments. In S. Sieloff Magnan (Ed.), *Mediating discourse online* (pp. 15-45). Amsterdam: John Benjamins.

Cobb, T. (2002). Web Vocabprofile. http://www.lextutor.ca/vp/

Cohen, A. D. (1996). Developing the ability to perform speech acts. *Studies in Second Language Acquisition, 18*(2), 253-267.

Cohen, A. D. (2005). Strategies for learning and performing L2 speech acts. *Intercultural Pragmatics, 2*(3), 275-301.

Cohen, A. D. (2008). Teaching and assessing L2 pragmatics: What can we expect from learners? *Language Teaching, 41*(2), 213–235.

Cohen, A. D., & Shively, R. (2007). Acquisition of requests and apologies in Spanish and French: Impact of study abroad and strategy-building Intervention. *The Modern Language Journal, 91*(2), 189-212.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2 ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cook, H. (2001). Why can't learners of JFL distinguish polite from impolite speech styles? In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 80-102). Cambridge: Cambridge University Press.

Cribbie, R. A., & Jamieson, J. (2004). Decreases in posttest variance and the measurement of change. *Methods of Psychological Research Online, 9*(1), 37-55. Retrieved from http://www.dgps.de/fachgruppen/methoden/mpr-online/issue22/mpr124_10.pdf

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*, 183–199.

Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing, 26*(3), 367–396.

de Saint Léger, D., & Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System, 37*(2), 269-285.

Derwing, T., Munro, M., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly, 34*, 592 - 603.

Derwing, T., Munro, M., Thomson, R., & Rossiter, M. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*(4), 533-557.

Dewaele, J.-M., & Pavlenko, A. (2003). Productivity and lexical diversity in native and non-native speech: A study of cross-cultural effects. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 120-141). Tonawanda, NY: Multilingual Matters.

Dewaele, J.-M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning, 58*(4), 911–960.

Dörnyei, Z. (2001). *Teaching and researching motivation*. Harlow: Longman.

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*, 3–32.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*, 42-68.

Elder, C., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 219-238). Amsterdam: John Benjamins.

Ellis, N. (2009). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 139-158). Oxford: Blackwell.

Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* (pp. 35-48). Washington DC: Georgetown University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. New York: Oxford University Press.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3-34). Amsterdam: John Benjamins.

Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Journal, 8*(3). Retrieved from http://www.asian-efl-journal.com/Sept_06_re.php

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*, 474-509.

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.

Feak, C., Reinhart, S., & Rohlck, T. (2009). *Academic interactions: Communicating on campus*. Ann Arbor, MI: University of Michigan Press.

Félix-Brasdefer, C. (2006). Teaching the negotiation of multi-turn speech acts: Using conversation-analytic tools to teach pragmatics in the classroom. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning, vol. 11* (pp. 165-197). Honolulu, HI: University of Hawai'i Press.

Félix-Brasdefer, C. (2010). Data collection methods in speech act performance: DCTs, role plays, and verbal reports. In A. Martínez-Flor & E. Usó-Juan (Eds.), *Speech act performance: Theoretical, empirical and methodological issues* (pp. 41–56). Amsterdam: John Benjamins.

Field, A. (2005). *Discovering statistics using SPSS* (2 ed.). London: Sage.

Field, A. (2009). *Discovering statistics using SPSS* (3 ed.). London: Sage.

Foster, P. (1996). Doing the task better: How planning time influences students' performance. In J. R. Willis & D. A. Willis (Eds.), *Challenge and change in language teaching* (pp. 126-135): Heinemann.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*, 299-323.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research, 3*(3), 215-247.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*, 354-375.

Freed, B. (Ed.). (1995). *Second language acquisition in a study abroad context*. Amsterdam: John Benjamins.

Gardner, R. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1-20). Hololulu: University of Hawai'i Press.

Garrett, N. (1991). Technology in the service of language learning: Trends and issues. *The Modern Language Journal, 75*, 74–101.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., & Houck, N. (1999). *Interlanguage refusals: A cross-cultural study of Japanese-English*. Berlin: Mouton de Gruyter.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., & Mackey, A. (2006). Input, interaction and output in SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175-199). Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., Mackey, A., Fernandez, M., & Alvarez-Torres, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*, 549–580.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3 ed.). New York: Routledge.

Gass, S. M., & Varonis, E. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition, 16*, 283-302.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and (+/- here-and-now): Effects on L2 oral production. In M. García Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 44-68). Tonawanda, NY: Multilingual Matters.

Guará-Tavares, M. G. (2008). *Pre-task planning, working memory capacity and L2 speech performance* (Unpublished doctoral dissertation). Universidade Federal de Santa Catarina, Brazil.

Harward, S., Allred, R., & Sudweeks, R. (1994). The effectiveness of four self-corrected spelling test methods. *Reading Psychology, 15*(4), 245-271.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.

Hinkle, D., Wiersma, W., & Jurs, S. (2003). *Applied statistics for the behavioral sciences*. Boston, MA: Houghton Mifflin.

Holmes, J. (2003). "I couldn't follow her story…": Ethnic differences in New Zealand narratives. In J. House, G. Kasper, & S. Ross (Eds.), *Misunderstanding in social life* (pp. 173-198). Harlow, UK: Pearson Education.

House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in Second Language Acquisition, 18*(2), 225-252.

House, J., & Kasper, G. (1981). Politeness markers in English and German. In F. Coulmas (Ed.), *Conversational routine: Explorations in standardized communication situations and prepatterned speech* (pp. 157-185). The Hague: Mouton.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*, 461-473.

Hubbard, P. (2004). Learner training for effective use of CALL. In S. Fotos & C. M. Browne (Eds.), *New perspectives on CALL for second language classrooms* (pp. 45-67). Mahwah, NJ: Lawrence Erlbaum Associates.

Hudson, T. (2001). Indicators for pragmatic instruction: Some quantitative tools. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 283-300). New York: Cambridge University Press.

Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics*. Honolulu: University of Hawai'i Press.

Hunt, K. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development, 35*(1), 1-61.

Ife, A., Vives Boix, G., & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics, 4*(1), 55-84.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics, 29*, 24 - 49.

Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 165–211). Amsterdam: John Benjamins.

Judd, E. L. (1999). Some issues in the teaching of pragmatic competence. In E. Hinkel (Ed.), *Culture in second language teaching and learning* (pp. 152-166). Cambridge: Cambridge University Press.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of English language learner proficiency in oral English. *The Modern Language Journal, 94*, 554-566.

Kang, S.-J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System, 33*, 277–292.

Kasper, G. (2004). Speech acts in (inter)action: Repeated questions. *Intercultural Pragmatics, 1*(1), 125-133.

Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning* (pp. 281-314). Honolulu, HI: University of Hawai'i Press.

Kasper, G., & Roever, C. (2005). Pragmatics in second language learning. In E. Hinkel (Ed.), *Handbook of research in second language learning and teaching* (pp. 317-334). Mahwah, NJ: Lawrence Erlbaum Associates.

Kasper, G., & Rose, K. (1999). Pragmatics and SLA. *Annual Review of Applied Linguistics, 19*, 81-104.

Kasper, G., & Rose, K. (2001). Pragmatics in language teaching. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 1-9). Cambridge: Cambridge University Press.

Kasper, G., & Rose, K. (2002). *Pragmatic development in a second language*. Oxford: Blackwell.

Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition, 18*(2), 149-169.

Keck, C. M., Iberri-Shea, G., Tracey-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based interaction and acquisition: A meta-analysis. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91-131). Amsterdam: John Benjamins.

Keselman, H., Huberty, C., Lix, L., Olejnik, S., Cribbie, R., Donahue, B., Kowalchuk, R., et al. (1998). Statistical practices of educational researchers: An analysis of the ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research, 68*, 350-386.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research, 12*, 211–234.

Kobayashi, M. (2003). The role of peer support in ESL students' accomplishment of oral academic tasks. *The Canadian Modern Language Review, 59*(3), 337-368.

Koike, D., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System, 33*(3), 481-501.

Kondo, S. (1997). The development of pragmatic competence by Japanese learners of English: Longitudinal study on interlanguage apologies. *Sophia Linguistica, 41*, 265-284.

Kondo, S. (2008). Effects on pragmatic development through awareness-raising instruction: Refusals by Japanese EFL learners. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 153-177). Tonawanda, NY: Multilingual Matters.

Kost, R. C. (2004). *An investigation of the effects of synchronous computer-mediated communication (CMC) on interlanguage development in beginning learners of German: Accuracy, proficiency, and communication strategies* (Unpublished doctoral dissertation). University of Arizona.

Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote metalinguistic awareness in the French immersion classroom? In R. Johnson & M. Swain

(Eds.), *Immersion education: International perspectives* (pp. 284–309). Cambridge: Cambridge University Press.

Kraemer, A. (2008). *Engaging the foreign language learner: Using hybrid instruction to bridge the language-literature gap* (Unpublished doctoral dissertation). Michigan State University.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. New York: Longman.

Kuriscak, L., & Luke, C. (2009). Language learner attitudes toward virtual worlds: An investigation of Second Life. In L. Lomicka & G. Lord (Eds.), *The next generation: Social networking and online collaboration in foreign language learning* (pp. 173-198). San Marcos, TX: CALICO Publications.

Lantolf, J. P., & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

Lazaraton, A., & Davis, L. (2008). A microanalytic perspective on discourse, proficiency, and identity in paired oral assessment. *Language Assessment Quarterly, 5*(4), 313–335.

Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the U.S. *Language Learning & Technology, 8*(1), 83-100. Retrieved from http://llt.msu.edu/vol8num1/pdf/lee.pdf

Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research, 8*, 55–81.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387-417.

Levy, M. (2009). Technologies in use for second language learning. *The Modern Language Journal, 93*(s1), 769–782.

Liddicoat, A., & Crozet, C. (2001). Acquiring French interactional norms through instruction. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 125-144). Cambridge: Cambridge University Press.

Liu, J. (2006). *Measuring interlanguage pragmatic knowledge of EFL learners*. Frankfurt am Main: Peter Lang.

LoCastro, V. (2003). *An introduction to pragmatics: Social action for language teachers*. Ann Arbor, MI: The University of Michigan Press.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press.

Lumley, T., Brown, A., & Zhang Wenxia, B. (2003, July). *Learner response to new TOEFL integrated reading/writing tasks.* Paper presented at the annual Language Testing Research Colloquium, The University of Reading, UK.

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research, 4*, 221-250.

Lynch, T., & Maclean, J. (2001). A case of exercising: Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 141–162). Harlow, UK: Pearson.

Machado, A. (2000). A Vygotskian approach to evaluation in foreign language learning contexts. *ELT Journal, 54*, 335–345.

MacIntyre, P. (2002). Motivation, anxiety and emotion in second language acquisition. In P. Robinson (Ed.), *Individual differences in second language acquisition* (pp. 45-68). Amsterdam: John Benjamins.

Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*, 557-587.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction and second language acquisition: A series of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal, 82*, 338-356.

Martínez-Flor, A., & Fukuya, Y. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System, 33*(3), 463-480.

Martínez-Flor, A., & Usó-Juan, E. (2006). A comprehensive pedagogical framework to develop pragmatics in the foreign language classroom: The 6Rs Approach. *Applied Language Learning, 16*(2), 39-64.

Martínez-Flor, A., & Usó-Juan, E. (2010). Pragmatics and speech act performance In A. Martínez-Flor & E. Usó-Juan (Eds.), *Speech act performance: Theoretical, empirical and methodological issues* (pp. 3-20). Amsterdam: John Benjamins.

Martínez-Flor, A., Usó-Juan, E., & Alcón Soler, E. (2006). Towards acquiring communicative competence through speaking. In E. Usó-Juan & A. Martínez-Flor (Eds.), *Current trends in the development and teaching of the four language skills* (pp. 139-158). Berlin: Walter de Gruyter.

Matsumura, S. (2001). Learning the rules for offering advice: A quantitative approach to second language socialization. *Language Learning, 51*, 635-679.

McCrae, R. R., & Costa, P. T. (2003). *Personality in adulthood: A five-factor theory perspective* (2 ed.). New York: Guilford Press.

McNamara, T. (1997). 'Interaction' in second language performance assessment: Whose performance? *Applied Linguistics, 18*, 446-466.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20*, 83-108.

Michel, M. C., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching, 45*, 241–259.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40-48.

Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research, 12*(1), 11-37.

Morton, H., & Jack, M. (2005). Scenario-based spoken interaction with virtual agents. *Computer Assisted Language Learning, 18*(3), 171 - 191.

Morton, H., & Jack, M. (2010). Speech interactive computer-assisted language learning: A cross-cultural evaluation. *Computer Assisted Language Learning, 23*(4), 295 - 319.

Németh, N., & Kormos, J. (2001). Pragmatic aspects of task-performance: The case of argumentation. *Language Teaching Research, 5*(3), 213-240.

Niezgoda, K., & Roever, C. (2001). Pragmatic and grammatical awareness. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 63-79). Cambridge: Cambridge University Press.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*, 555-578.

Ohta, A. S. (2001). Peer interactive tasks and assisted performance in classroom language learning. In A. S. Ohta (Ed.), *Second language acquisition processes in the classroom: Learning Japanese* (pp. 73-128). Mahwah, NJ: Lawrence Erlbaum Associates.

Olshtain, E., & Cohen, A. D. (1991). Teaching speech act behavior to nonnative speakers. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 154-165). Boston, MA: Heinle & Heinle.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*, 109–148.

Pawley, A., & Syder, F. (2000). The one clause at a time hypothesis. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 163–191). Ann Arbor: University of Michigan Press.

Pedhazur, E. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3 ed.). Fort Worth, TX: Harcourt Brace College Publishers.

Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*, 101–143.

Porter, P. (1986). How learners talk to each other: Input and interaction in task-centered discussions. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 200-224). Rowley, MA: Newbury House.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45*, 99-140.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*, 533-554.

Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing, 23*(2), 229-256.

Roever, C. (2008). Rater, item, and candidate effects in discourse completion tests: A FACETS approach. In A. Martínez-Flor & E. Alcón Soler (Eds.), *Investigating pragmatics in foreign language learning, teaching, and testing* (pp. 249–266). Clevedon, UK: Multilingual Matters.

Roever, C. (2009). Teaching and testing pragmatics. In M. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 560-577). Oxford: Wiley-Blackwell.

Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing, 28*(2), 1-19.

Rose, K. (2000). An exploratory cross-sectional study of interlanguage pragmatic development. *Studies in Second Language Acquisition, 22*, 27-67.

Rose, K. (2005). On the effects of instruction in second language pragmatics. *System, 33*, 385-399.

Rose, K., & Ng, C. (2001). Inductive and deductive teaching of compliments and compliment responses. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 145-170). Cambridge: Cambridge University Press.

Rossiter, M. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review, 65*(3), 395-412.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Amsterdam: John Benjamins.

Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111-141). Amsterdam: John Benjamins.

Sauro, S., & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics, 31*(4), 554-577.

Savignon, S. (1997). *Communicative competence: Theory and classroom practice*. New York: McGraw-Hill.

Savignon, S. (2001). Communicative language teaching in the twenty-first century. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 13-28). Boston: Heinle & Heinle.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness. In R. Schmidt (Ed.), *Attention and awareness in foreign language teaching and learning* (pp. 1-64). Honolulu: University of Hawai'i at Mānoa.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Searle, J. (1969). *Speech acts*. Cambridge: Cambridge University Press.

Segalowitz, N. (2003). Automaticity and second language acquisition. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382-408). Oxford: Blackwell Publishers.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.

Shardakova, M. (2005). Intercultural pragmatics in the speech of American L2 learners of Russian: Apologies offered by Americans in Russian. *Intercultural Pragmatics, 2*(4), 423-451.

Shea, M. (2011, March). *Clause or phrase? Comparing syntactic complexity measures.* Paper presented at the annual meeting of the American Association of Applied Linguistics, Chicago, IL.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*, 38-62.

Skehan, P. (2002). A non-marginal role for tasks. *ELT Journal, 56*, 289–295.

Skehan, P. (2003). Focus on form, tasks, and technology. *Computer Assisted Language Learning, 16*, 391-411.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*, 510-532.

Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research, 1*, 185–211.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task-performance in a second language* (pp. 193-216). Amsterdam: John Benjamins.

Sundqvist, P. (2009). *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary* (Unpublished doctoral dissertation). Karlstad University.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Siedlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford, UK: Oxford University Press.

Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing, 18*, 275-302.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471-484). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step toward second language learning. *Applied Linguistics, 16*, 371–391.

Sykes, J. M. (2005). Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO Journal, 22*, 399-431.

Sykes, J. M. (2008). *A dynamic approach to social interaction: SCMC, synthetic immersive environments & Spanish pragmatics* (Unpublished doctoral dissertation). University of Minnesota.

Sykes, J. M., Oskoz, A., & Thorne, S. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal, 25*, 528-546.

Taguchi, N. (2006). Analysis of appropriateness in a speech act of request in L2 English. *Pragmatics*, 513-533.

Taguchi, N. (2007). Task difficulty in oral speech act production. *Applied Linguistics, 28*, 113-135.

Takahashi, S. (2001). The role of input enhancement in developing interlanguage pragmatic competence. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 171-199). New York: Cambridge University Press.

Takimoto, M. (2006). The effects of explicit feedback and form-meaning processing on the development of pragmatic proficiency in consciousness-raising tasks. *System, 34*(4), 601-614.

Takimoto, M. (2009). Exploring the effects of input-based treatment and test on the development of learners' pragmatic proficiency. *Journal of Pragmatics, 41*(5), 1029-1046.

Tateyama, Y. (2001). Explicit and implicit teaching of pragmatic routines. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 200-222). Cambridge: Cambridge University Press.

Tateyama, Y., Kasper, G., Mui, L., Tay, H., & Thananart, O. (1997). Explicit and implicit teaching of pragmatic routines. In L. F. Bouton (Ed.), *Pragmatics and language learning: Monograph series vol. 8* (pp. 163-178). Urbana-Champaign, IL: University of Illinois, Urbana-Champaign.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*, 439–473.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-276). Amsterdam: John Benjamins.

Thomson, R., & Isaacs, T. (2010, March). *The real me vs. artificial her: The effect of grammatical person, topic familiarity and task type on L2 oral performance.* Paper presented at the annual meeting of the American Association for Applied Linguistics, Atlanta, GA.

Thorne, S. (2003). Artifacts and cultures-of-use in intercultural communication. *Language Learning & Technology, 7*(2), 38-67.

Trosborg, A. (1995). *Interlanguage pragmatics: Requests, complaints and apologies.* Berlin: Mouton de Gruyter.

Usó-Juan, E. (2008a). A pragmatic-focused evaluation of requests and their modification devices in textbook conversations. In E. Alcón Soler (Ed.), *Learning how to request in an instructed language learning context* (pp. 65-90). Bern: Peter Lang.

Usó-Juan, E. (2008b). The presentation and practice of the communicative act of requesting in textbooks: Focusing on modifiers. In E. Alcón Soler & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 223-244). Dordrecht: Springer.

Vellenga, H. (2004). Learning pragmatics from ESL and EFL textbooks: How likely? *TESL-EJ, 8*(2), 1-18.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17*(1), 65-83.

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*, 121–142.

Weissheimer, J. (2007). *Working memory capacity and the development of L2 speech production: An exploratory study* (Unpublished doctoral dissertation). Universidade Federal de Santa Catarina, Florianopolis.

Wendel, J. (1997). *Planning and second language narrative production* (Unpublished doctoral dissertation). Temple University, Japan.

Wesche, M. B., & Skehan, P. (2002). Communicative, task-based, and content-based language instruction. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 207-288). New York: Oxford University Press.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing, 14*, 85–106.

Wilkinson, S. (2002). The omnipresent classroom during summer study abroad: American students in conversation with their French hosts. *The Modern Language Journal, 86*, 157-173.

Williams, J. (1999). Learner-generated attention to form. *Language Learning, 51*, 303–346.

Willis, J. R. (2004). Perspectives on task-based instruction: Understanding our practices, acknowledging different practitioners. In B. Leaver & J. R. Willis (Eds.), *Task-based instruction in foreign language education: Practices and programs* (pp. 3-44). Washington, DC: Georgetown University Press.

Winke, P., & Goertler, S. (2008). Did we forget someone? Students' computer access and literacy for CALL. *CALICO Journal, 25*(3), 482-509.

Winke, P., & Teng, C. (2010). Using task-based pragmatics tutorials while studying abroad in China. *Intercultural Pragmatics, 7*(2), 363-399.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu, HI: University of Hawai'i Press.

Woodrow, L. (2006). Anxiety and Speaking English as a Second Language. *RELC Journal, 37*(3), 308-328.

Yoshimi, D. (2001). Explicit instruction and JFL learners' use of interactional discourse markers. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 223-244). Cambridge: Cambridge University Press.

Young, L. W. L. (1994). *Crosstalk and culture in Sino-American communication*. Cambridge: Cambridge University Press.

Yu, M.-C. (2008). Teaching and learning sociolinguistic skills in university EFL classes in Taiwan. *TESOL Quarterly, 42*(1), 31-53.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*, 1–27.

Zyzik, E., & Polio, C. (2008). Incidental focus on form in Spanish literature courses. *The Modern Language Journal, 92*(1), 50-73.