2007

Instruction and Developing Second Language Pragmatic Competence: An Investigation into the Efficacy of Output

Justin E. (Justin Edward) Jernigan
THE FLORIDA STATE UNIVERSITY

COLLEGE OF EDUCATION

INSTRUCTION AND DEVELOPING SECOND LANGUAGE PRAGMATIC
COMPETENCE: AN INVESTIGATION INTO THE EFFICACY OF OUTPUT

By

JUSTIN E. JERNIGAN

A Dissertation submitted to the
Department of Middle and Secondary Education
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 2007
The members of the Committee approve the dissertation of Justin E. Jernigan defended on January 19, 2007.

________________________(signed)
Deborah J. Hasson
Professor Directing Dissertation

________________________(signed)
Michael J. Leeser
Outside Committee Member

________________________(signed)
Patrick C. Kennell
Committee Member

________________________(signed)
Sharilyn Steadman
Committee Member

Approved:

________________________
Pamela Sissi Carroll, Chair, Department of Middle and Secondary Education

The Office of Graduate Studies has verified and approved the above named committee members.
For Vien: None of this would have been possible without your sacrifices and support. I love you.
I wish to express my sincere gratitude to the esteemed members of my dissertation committee: Dr. Deborah J. Hasson, my major professor and chair of the committee, who provided essential stability, direction, and encouragement, as well as a keen eye for detail throughout the process of writing and revising the dissertation; Dr. Michael J. Leeser, who was always available for discussion and offered invaluable advice and guidance concerning the design of the study, the analysis of the data, and formatting issues; Dr. Patrick C. Kennell, who, in his role as both committee member and Director of the Center for Intensive English Studies, facilitated the carrying out of this research in many ways and encouraged me as a researcher, teacher, and family man; and Dr. Sharilyn Steadman, who offered her expertise in reading the dissertation closely and providing important feedback. I am proud to call the members of my committee not only mentors, but also friends.

I also wish to thank the following, each of whom contributed to the planning and completion of the dissertation in important ways: Dr. N. Eleni Pappamihiel, who supervised my prospectus and served as a motivating mentor and guide during the initial planning and writing of the dissertation; Dr. Elizabeth Platt, who had a vital role in shaping my early research and was a source of useful ideas on several occasions; Dr. Bill VanPatten, who offered creative insight and valuable suggestions concerning the design of the study and my model of L2 pragmatic acquisition; Dr. Akihito Kamata, who provided me with needed insight into the statistical analyses conducted in the study; and Ramin Yazdanpanah, whose help and cooperation in the production of the video materials and data collection was indispensable. I would also like to express my thanks to the other instructional staff of the Center for Intensive English Studies, who gave of their time and talents when called upon during the preparation of materials and gathering of data.

I acknowledge the important contributions of all of those mentioned above. Of course, any shortcomings or limitations in the study are my own.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xi</td>
</tr>
<tr>
<td><strong>CHAPTER 1</strong></td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION AND BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>3</td>
</tr>
<tr>
<td>Developing L2 Pragmatic Competence in SLA</td>
<td>3</td>
</tr>
<tr>
<td>The Focus of the Present Study</td>
<td>7</td>
</tr>
<tr>
<td>The Theoretical Significance of the Present Study</td>
<td>11</td>
</tr>
<tr>
<td>Research Questions</td>
<td>12</td>
</tr>
<tr>
<td>Summary</td>
<td>13</td>
</tr>
<tr>
<td>Outline of the Dissertation</td>
<td>14</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>14</td>
</tr>
<tr>
<td><strong>CHAPTER 2</strong></td>
<td></td>
</tr>
<tr>
<td>LITERATURE REVIEW AND MOTIVATION OF THE VARIABLES</td>
<td>17</td>
</tr>
<tr>
<td>Defining Pragmatic Competence: Key Constructs</td>
<td>18</td>
</tr>
<tr>
<td>Facilitating L2 Pragmatic Competence: Key Questions</td>
<td>25</td>
</tr>
<tr>
<td>L2 Pragmatic Speech Act Development</td>
<td>42</td>
</tr>
<tr>
<td>Assessing Developing L2 Pragmatic Performance</td>
<td>47</td>
</tr>
<tr>
<td>Summary</td>
<td>48</td>
</tr>
<tr>
<td><strong>CHAPTER 3</strong></td>
<td></td>
</tr>
<tr>
<td>RESEARCH DESIGN AND METHODOLOGY</td>
<td>50</td>
</tr>
<tr>
<td>The Pilot Study</td>
<td>50</td>
</tr>
<tr>
<td>The Principal Research Study</td>
<td>56</td>
</tr>
<tr>
<td>Research Design</td>
<td>57</td>
</tr>
<tr>
<td>Participants</td>
<td>60</td>
</tr>
<tr>
<td>Materials</td>
<td>64</td>
</tr>
<tr>
<td>Procedure</td>
<td>68</td>
</tr>
</tbody>
</table>
### Scoring and Analyses

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
</tr>
</tbody>
</table>

### CHAPTER 4

### RESULTS

<table>
<thead>
<tr>
<th>Research Questions Revisited</th>
<th>83</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Reconstruction</td>
<td>84</td>
</tr>
<tr>
<td>Statistical Tests</td>
<td>85</td>
</tr>
<tr>
<td>The Pragmatic Acceptability Judgment Task</td>
<td>88</td>
</tr>
<tr>
<td>The Written Discourse Completion Task</td>
<td>94</td>
</tr>
<tr>
<td>The Oral Discourse Completion Task</td>
<td>100</td>
</tr>
<tr>
<td>The DCT Rating Rubric</td>
<td>107</td>
</tr>
<tr>
<td>Summary</td>
<td>108</td>
</tr>
</tbody>
</table>

### CHAPTER 5

### DISCUSSION AND CONCLUSIONS

<table>
<thead>
<tr>
<th>Overview</th>
<th>111</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constructs: Output and Developing L2 Pragmatic Competence</td>
<td>112</td>
</tr>
<tr>
<td>Research Question 1</td>
<td>113</td>
</tr>
<tr>
<td>Research Question 2</td>
<td>117</td>
</tr>
<tr>
<td>Research Question 3</td>
<td>121</td>
</tr>
<tr>
<td>Summary of Questions and Findings</td>
<td>122</td>
</tr>
<tr>
<td>Limitations</td>
<td>122</td>
</tr>
<tr>
<td>Future Directions for Research</td>
<td>126</td>
</tr>
<tr>
<td>Recommendations for Acquisitional Pragmatics and ESL Practitioners</td>
<td>129</td>
</tr>
<tr>
<td>Concluding Thoughts</td>
<td>133</td>
</tr>
</tbody>
</table>

### APPENDIX A: Pragmatic Acceptability Judgment Task-Pretest

<p>| APPENDIX B: Pragmatic Acceptability Judgment Task-Posttest | 134 |
| APPENDIX C: Written Discourse Completion Task-Pretest    | 138 |
| APPENDIX D: Written Discourse Completion Task-Posttest   | 142 |
| APPENDIX E: Oral Discourse Completion Task-Pretest       | 145 |
| APPENDIX F: Oral Discourse Completion Task-Posttest      | 148 |
| APPENDIX G: DCT Rater Training                         | 151 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Grice’s (1975) Conversational Maxims</td>
<td>22</td>
</tr>
<tr>
<td>2.</td>
<td>Searle’s Classification of Speech Acts</td>
<td>24</td>
</tr>
<tr>
<td>3.</td>
<td>Studies on Instruction and Developing L2 Pragmatics</td>
<td>33</td>
</tr>
<tr>
<td>4.</td>
<td>Proposed Effects of Output on Developing L2 Pragmatics</td>
<td>38</td>
</tr>
<tr>
<td>5.</td>
<td>Studies on Output and Developing L2 Pragmatics</td>
<td>41</td>
</tr>
<tr>
<td>6.</td>
<td>Ohta’s (2001) Six Stages of JFL Alignment Expression Acquisition</td>
<td>43</td>
</tr>
<tr>
<td>8.</td>
<td>Mean Pilot DCT Scores by Output Group Condition</td>
<td>54</td>
</tr>
<tr>
<td>9.</td>
<td>Summary of Research Questions</td>
<td>56</td>
</tr>
<tr>
<td>10.</td>
<td>Summary of Variables in the Study</td>
<td>58</td>
</tr>
<tr>
<td>11.</td>
<td>Summary of Speech Acts Tested</td>
<td>60</td>
</tr>
<tr>
<td>12.</td>
<td>+ Output Group Study Participants</td>
<td>62</td>
</tr>
<tr>
<td>13.</td>
<td>- Output Group Study Participants</td>
<td>63</td>
</tr>
<tr>
<td>14.</td>
<td>Daily Schedule for Instructional Treatment and Data Collection</td>
<td>73</td>
</tr>
<tr>
<td>15.</td>
<td>Scoring Scale for Rating DCT Tasks</td>
<td>78</td>
</tr>
<tr>
<td>16.</td>
<td>Examples Corresponding to Each DCT Rating Level</td>
<td>79</td>
</tr>
<tr>
<td>17.</td>
<td>Text Reconstruction Examples</td>
<td>85</td>
</tr>
<tr>
<td>18.</td>
<td>PAJT Descriptive Statistics</td>
<td>88</td>
</tr>
<tr>
<td>19.</td>
<td>PAJT ANOVA Table</td>
<td>89</td>
</tr>
<tr>
<td>20.</td>
<td>PAJT t-test Table</td>
<td>89</td>
</tr>
<tr>
<td>21.</td>
<td>PAJT Effect Sizes</td>
<td>90</td>
</tr>
<tr>
<td>22.</td>
<td>PAJT Directives Descriptive Statistics</td>
<td>91</td>
</tr>
<tr>
<td>23.</td>
<td>PAJT Directives ANOVA Table</td>
<td>91</td>
</tr>
<tr>
<td>24.</td>
<td>PAJT Directives t-test Table</td>
<td>92</td>
</tr>
<tr>
<td>25.</td>
<td>PAJT Directives Effect Sizes</td>
<td>92</td>
</tr>
<tr>
<td>26.</td>
<td>PAJT Expressives Descriptive Statistics</td>
<td>93</td>
</tr>
<tr>
<td>27.</td>
<td>PAJT Expressives ANOVA Table</td>
<td>93</td>
</tr>
<tr>
<td>28.</td>
<td>PAJT Expressives Effect Sizes</td>
<td>94</td>
</tr>
<tr>
<td>29.</td>
<td>WDCT Descriptive Statistics</td>
<td>95</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Model of Developing Interlanguage Pragmatic Competence Acquisition ............... 9
3. Sample Pilot Test DCT Item ........................................................................ 51
4. Sample Pilot Test Comprehension Question ................................................. 52
5. Overview of Research Design ...................................................................... 57
6. Sample Pragmatic Video Vignette ................................................................ 65
7. Text Reconstruction and Comprehension Instructions .................................... 71
8. Sample Speaker Intention-focused PAJT Item .............................................. 74
9. Sample Speech Act-focused PAJT Item ......................................................... 75
10. PAJT and WDCT/ODCT Tasks Related to Speech Act Types ...................... 120
ABSTRACT

The primary purpose of the present study was to investigate the efficacy of output in the context of instruction on developing second language (L2) pragmatic competence. The study focused on the ways in which opportunities for output raise adult ESL learners’ awareness of pragmalinguistic forms, thus facilitating the acquisition of those forms. The setting of the study was an intensive English program (IEP) in the United States, and the research was conducted within the broad framework of Swain’s (1985, 1995, 2005) Output Hypothesis.

In the area of developing L2 pragmatics research, efforts to understand a possible role for Output have been extremely limited. In fact, developing L2 pragmatics research that focuses on acquisitional processes is a young subdiscipline, having emerged on the scene of pragmatics research only within the last two decades (Bardovi-Harlig, 1999; Kasper & Roever, 2005). Also, the focus of much developing L2 pragmatics research has been the effects of instruction on learners’ developing L2 pragmatics (e.g., Ohta, 2001; Rose & Ng, 2001; Takahashi, 2005b). Although such efforts have implications for the present study, they have done little to make clear the possible contributions of Output to learners’ developing L2 pragmatics. Studies with greater implications for the exploration of Output (e.g., Bardovi-Harlig & Salsbury, 2004; Liddicoat & Crozet, 2001; Martínez-Flor & Fukuya, 2005) are explored in the review of literature, but are shown to have many limitations as well. For the most part, the construct of Output as a causative variable in developing L2 pragmatic competence has gone unexplored in any meaningful way.

The research questions addressed in the present study were:

1. Is there an effect for Output in the context of video-based pragmatic instruction (including output-focused tasks) on the developing L2 pragmatic competence of adult ESL learners, as evidenced by the acquisition of pragmalinguistic forms (pragmatic speech acts) in appropriate contexts?

   1a. Is there an effect for Output on pragmatic perception on a pragmatic acceptability judgment task (PAJT; Garcia, 2004) from pretest to posttest?

   1b. Is there an effect for Output on written pragmalinguistic production on a written discourse completion task (WDCT) from pretest to posttest?

   1c. Is there an effect for Output on oral pragmalinguistic production on an oral
discourse completion task (ODCT) from pretest to posttest?

2. Is there a differential effect of Output on the two different types of pragmatic speech acts of directives and expressives?

   2a. Is there a differential effect for Output on the perception of directives and expressives on a pragmatic acceptability judgment task (PAJT) from pretest to posttest?

   2b. Is there a differential effect for Output on the production of directives and expressives on a written discourse completion task (WDCT) from pretest to posttest?

   2c. Is there a differential effect for Output on the production of directives and expressives on an oral discourse completion task (ODCT) from pretest to posttest?

3. Is the proposed rubric a reliable instrument for the assessment of the pragmatic acceptability of the responses of adult ESL learners on discourse completion tasks (DCTs)?

   The variables tested in the present study included Output (two groups: + Output and – Output); time (pretest to posttest); learners’ perceptions of pragmalinguistic forms (pretest-posttest); and the production of pragmalinguistic forms in oral and written modes (pretest-posttest). In addition, following Yamashita (1996), the rubric proposed in the present study for rating the participants’ responses in terms of pragmatic acceptability was tested for reliability. This examination of the rubric itself allowed the first two research questions to be addressed with more confidence with respect to the validity and reliability of the measures.

   The findings were that, with respect to Research Question 1, the + Output instructional treatment had a significant effect on the + Output group participants’ pretest-posttest performance on the PAJT. Also, a large effect size was calculated for the + Output treatment group. No significant effect was observed for the – Output group, and the effect size calculated was below medium in size. No significant effects were identified for either the + Output or – Output group on the WDCT task, although a relatively large effect size was calculated for the + Output group. Finally, on the ODCT task, there was a significant effect observed for the – Output treatment group and not for the + Output group, although both groups had relatively large effect sizes associated with their ODCT pretest-posttest performances.
With respect to Research Question 2, the findings of the study revealed that there were some significant differences in the way the instructional treatment impacted the learners’ perception and production of the two different types of speech act tested (i.e., directives and expressives). For the PAJT task, the + Output treatment had a significant effect on directives, but not on expressives. On the WDCT task, the learners in the + Output group demonstrated significant change from pretest to posttest for the expressive speech act items. Large effect sizes were also calculated for both instructional treatments on the expressives. There were no significant effects identified on the directive items. Finally, on the ODCT task, there were significant pretest-posttest effects with large effect sizes calculated for both the + Output and – Output instructional treatments with respect to the expressive items.

Research Question 3 addressed the reliability and internal consistency of the 4-level rating rubric used to assess the acceptability of the participants’ responses on the DCT instruments. The findings indicated acceptably high inter-rater reliability levels for three of the four instruments tested (the WDCT pretest and posttest and the ODCT pretest), whereas the consistency of the rating scores on the ODCT posttest were just below what is normally desired in social science research.

The limitations facing the study are acknowledged. First, the ability to judge pragmatic responses is a subjective, human endeavor, and so is subject to error. In order to control for this potential source of error, multiple raters were employed for the scoring of responses, and consistency and reliability promoted. Also, the ability to isolate Output as a causative variable is no easy task. The study followed previous research (e.g., Beebe, Takahashi, & Uliss-Weltz, 1990; Izumi, 2002) in research and task design as far as possible in order to promote the validity and reliability of all aspects of the study.

Another possible contribution of the study to the field of developing L2 pragmatics research, suggested in the second research question, is the addition of support to the growing numbers of studies that have proposed developmental stages for L2 pragmatics (e.g., Bardovi-Harlig & Salsbury, 2004; Ohta, 2001). The targeting of two categories of speech act in the present study represented an effort to understand whether different pragmalinguistic forms respond differently to the output-focused video pragmatic instruction featured in the treatment. If a developmental stage exists for pragmalinguistic forms, it is reasonable to expect that some forms might be acquired earlier than others.
The acquisition of second language (L2) pragmatic competence by adult language learners is a daunting task by any measure. The difficulties attached to L2 pragmatic competence acquisition have to do with the complex nature of the process itself. In order to acquire pragmatic competence, learners must develop in terms of not only linguistic competence, but also in terms of sociocultural awareness, attaining a useful understanding of how language functions in social and cultural contexts (Kasper & Roever, 2005). In addition, it can be difficult for learners to grasp with precision what is pragmatically acceptable in different cultural and sub-cultural contexts (Barron, 2005). For example, Davies and Tyler (2005) reported on the pragmatic difficulties faced by a Korean international teaching assistant (ITA) when confronting a U.S. student in the context of the American university classroom sub-culture. Given the challenges presented by L2 pragmatic acquisition, greater understanding of the process and those factors that may contribute to pragmatic acquisition in additional languages is a worthwhile goal, and should help to illuminate further some aspects of the overall process of second language acquisition (SLA) as well.

The study of the intersection of pragmatics and linguistics, referred to specifically as *pragmalinguistics* in such early and theoretically diverse works on pragmatics as Mey (1979) and Thomas (1983), gave rise to questions of how the general language faculty and the pragmatic system correspond. Ochs (1979) and Schieffelin (1979), for example, discussed the contribution of ethnographic observation of child language use to the understanding of developmental first language pragmatics in an important volume that they co-edited. Logically, researchers focusing on second language acquisition and those exploring issues connected with pragmatics began exploring the question of whether pragmatic competence in a second language was acquired in the same way as general linguistic competence. Efforts that explored interlanguage pragmatics more specifically (e.g., Kasper & Schmidt, 1996) emerged as second language pragmatics research began to take shape as a research discipline in its own right within the larger field of SLA research. In fact, the stated goal of much developing L2 pragmatics research since that time...
has been to push research efforts in this area to address questions of theoretical interest to general SLA theory to a greater extent (Bardovi-Harlig, 1999; Kasper, 2001; Kasper & Roever, 2005).

Much recent research in the area of L2 pragmatic development has focused on input, individual differences, and noticing (e.g., Kasper & Rose, 2002; Takahashi, 2005a) as well as pragmatic transfer from a learner’s first language (Al-Issa, 2003; Byon, 2004; Rose, 2000; Yoon, 1991) and classroom instructional practices (e.g., Davies, 2004; Rose, 2005; Rose & Kasper, 2001; Yates, 2004) as they affect the acquisition of pragmatic competence in L2. Another aspect of developmental pragmatics research that has received much attention recently is the type of data that are best suited for pragmatics studies. Bardovi-Harlig and Hartford (2005) have pointed out the desirability of conducting research in interlanguage pragmatics that recognizes the tension between “highly controlled production tasks that yield comparable language samples and… the investigation of authentic discourse” (p. 1). Ohta (2001, 2005), among others, has charted a course in interlanguage and developing L2 pragmatics research that relies almost exclusively on naturally occurring classroom data and naturalistic observational research.

Among the more commonly addressed questions in the L2 developmental pragmatics literature are: (a) How does pragmatic competence relate to other aspects of language competence? (b) How does a second language learner’s first language and culture impact her acquisition of pragmatic competence in L2? (c) Is pragmatic competence in a second language teachable (i.e., is there an effect for instruction on L2 pragmatic competence acquisition)? (d) If so, are different methods of instruction more effective than others in terms of facilitating the acquisition of L2 pragmatic competence? Kasper and Rose (2002) have posited three basic types of questions that studies on the effect of instruction on pragmatics seek to answer:

1. Is the targeted pragmatic feature teachable at all?
2. Is instruction in the targeted feature more effective than no instruction?
3. Are different teaching approaches differentially effective? (p. 249).

The present study contributes specifically to the existing research literature that investigates the effects of specific instructional approaches on developing L2 pragmatic competence, and thus is relevant for Kasper and Rose’s first and third questions. The research questions, discussed later in this chapter, do not directly contrast instruction with no instruction, but certainly have implications for the overall teachability of the pragmatic speech acts considered.
Statement of the Problem

Relative to the aspects of pragmatics research presented above, there has been little emphasis placed on the possible impact of Output on pragmatic development in the literature. Although the role of instruction as it relates to L2 pragmatic development is being addressed in a growing number of studies, much of the research has tended to center around the questions of teachability of pragmatic features, the relative effectiveness of instruction as compared to natural exposure, and the possibility of different effects for different instructional methods (Kasper & Roever, 2005; Kasper & Rose, 2002; Rose & Kasper, 2001). A possible role for Output in instruction targeting L2 pragmatic forms has been mentioned (Martínez-Flor & Fukuya, 2005), but has not been discussed to any great extent in the literature to date. The present research effort is an attempt to address this relatively underexplored aspect of developmental pragmatics.

This chapter first defines the concept of developing L2 pragmatic competence, situating it within the SLA process by means of a discussion of three key relationships: (a) SLA and communicative competence, (b) SLA and pragmatic competence, and (c) instruction and developing L2 pragmatic competence. Each of the three relationships reveals progressively more about the central construct of interest in the study: developing L2 pragmatic competence. Following the discussion of the three key relationships, the focus and theoretical significance of the present study is explained. Chapter 1 concludes with an outline of the dissertation and definitions of key terms used in the dissertation.

Developing L2 Pragmatic Competence in SLA

Efforts to understand the acquisition of pragmatics have not provided a coherent picture of the process, and calls for more systematic investigations that encompass more of the complexity of the acquisition of L2 pragmatic competence have persisted from the early years of interlanguage pragmatic research (e.g., Thomas, 1983) to more recent efforts (e.g., Hassall, 2004). In order to understand developing L2 pragmatics more fully, it is helpful to situate the process within the larger second language acquisition (SLA) process (Kasper & Schmidt, 1996). Here, the goal of locating developing L2 pragmatic competence within SLA is addressed through
an examination of the connections between SLA and communicative competence, SLA and pragmatic competence, and instruction and developing L2 pragmatic competence.

**SLA and Communicative Competence**

From the mainstream SLA perspective, Long (1990) discussed communicative competence as involving speakers’ abilities to encode and decode meaning from utterances according to the situation in which a given utterance occurs. Second language learners, Long argued, have a need to use the new language for communicative purposes, making developing communicative competence an integral part of successful SLA, along with linguistic development. In Swain’s (1985) study of French immersion students in Canada, the same construct, communicative competence, was conceived of as comprising grammatical, discourse, and sociolinguistic competencies. Of these three areas of language competence, mainstream SLA research has tended to examine the acquisition of grammatical and discourse competence more closely than sociolinguistic competence, with sociolinguistic concerns often explored in terms of attempts to define the context within which SLA processes unfold (Siegel, 2003), rather than in terms of understanding the ways in which individual learners are able to deal with the constraints that using language in those social settings imposes (Kasper & Rose, 2002).

A learner in the process of acquiring a second language must be able to process linguistic input successfully, to be sure. However, without the ability to understand the context in which an utterance is produced and the ways in which that context affects the discourse properties and sociolinguistic impact of the utterance, the learner cannot be said to have acquired the new language. In that sense, the learner’s ability to successfully develop overall L2 communicative competence, and not only linguistic (grammatical) competence, is a major feature of SLA. In further support of the notion that grammatical competence without accompanying sociolinguistic competence and strategies is insufficient to account for successful SLA, Bosco, Bucciarelli, and Bara (2004) noted that the very meaning of many speech acts depends on the context. They cite the example of a request that is issued in a context in which there are obvious obstacles to the hearer’s carrying out the request. In such a case, the apparently direct speech act of request may be interpreted instead as an indirect speech act. Cummings (2005) has stated that, “no definition of pragmatics would be complete in the absence of some mention of context” (p. 4). Based on her assertion, the investigation of developing L2 pragmatics appears to be an integral part of attaining a more complete understanding of the SLA process.
Ellis (1992) also offered evidence for the importance of the sociolinguistic component of communicative competence in a study of the requests produced by two child ESL learners over a period of almost two years. He found that the learners’ development in the use of requests was curtailed. Ellis proposed that the reason for the students’ inability to progress beyond a certain level in the production of this pragmatic speech act was that their sociolinguistic competence was not sufficiently developed. The study is revelatory of an important phenomenon in the L2 pragmatic acquisition process: learners’ progress in acquiring pragmatic competence is enhanced when all aspects of communicative competence are developed. Without development in the area of grammatical accuracy, pragmalinguistic production suffers. Likewise, without attention to sociolinguistic aspects of learning the new language and culture, the sociopragmatic domain of pragmatic competence is likely to fall behind in terms of development. Based on Lyster’s (1994) definition of sociolinguistic competence as “the ability to recognize and produce socially appropriate language in context” (p. 263), it is clear that the construct addressed in Ellis’ study has direct relevance to an understanding of L2 pragmatic competence, although developing L2 pragmatics may be distinguished as having a more specific focus on the perspective and intentions of the language learner during communication (Mey, 2001).

**SLA and Pragmatic Competence**

Pragmatic competence may be regarded as a component of communicative competence, but in discussing the place of pragmatic competence within the larger SLA process, the importance of the situation, or social and cultural context, in which communication occurs, alluded to in the discussion of communicative competence above, becomes even more prominent. In fact, one of the most important skills associated with pragmatic competence is the ability to recognize the appropriateness of an utterance within a given context and to choose one possible form over another based on that understanding (Kasper & Rose, 2002). For that reason, it may be claimed that without the ability to distinguish among different contexts in terms of their unique requirements of formality, politeness, and specialized lexicons, etc., a learner cannot be said to have acquired the pragmatic norms of the L2. In such a case, successful SLA could not be said to have occurred, regardless of the learner’s apparent gains in terms of grammatical and lexical awareness in the target language.

As a learner moves through the process of acquiring a second language, she is acquiring much more than the surface forms of the language or even the semantic and syntactic awareness
that accompany successful SLA. The learner is also internalizing the pragmatic constraints that apply in a variety of sociolinguistic contexts and developing pragmatic strategies by which to make use of the context to convey intended meanings. In the context of the present study, what is crucial in the discussion of how learners acquire both the L2 and the pragmatic system of the second language and culture is that the focus remain on cognitive acquisitional processes. In the field of developing L2 pragmatics research, maintaining an acquisitional focus in research can be challenging (Boxer, 2004). The challenge is posed by the fact that an understanding of social context is inseparable from any definition of pragmatic competence (Cummings, 2005). Thus, for the acquisition-focused L2 pragmatics researcher, there is necessarily a tension between the need to acknowledge the inherent social and cultural situatedness of pragmatic performance and to focus on the cognitively oriented acquisitional process of L2 pragmatic development.

Distinguishing pragmatic from morphosyntactic and lexical knowledge in a second language is likewise not a simple task. Kasper and Roever (2005) argued that, though L2 learners’ pragmatic abilities appear to develop in line with their overall proficiency, the correlation is not always linear, and does not appear to obtain in all domains of language use. Because of the interaction between the grammatical component and the interlanguage pragmatic component in a learner’s developing system during the development of overall communicative competence, it is probably futile to attempt to pull apart linguistic and pragmatic competence in L2 developmental pragmatics research. Rather, it is advisable to target aspects of learners’ L2 production that clearly indicate development in the area of pragmatic awareness and mastery. In this way, the role of grammatical competence and semantic lexical knowledge is acknowledged and accounted for, but is not the focus of the research.

**Instruction and Developing L2 Pragmatic Competence**

Much research over the past several years has focused on the effects of various aspects of L2 instruction on learners’ developing pragmatic competence (e.g., Liddicoat & Crozet, 2001; Rose & Ng, 2001; Tateyama, 2001). Following Rose (2005), among the key questions that have been addressed in many of the instruction and pragmatic development studies have been: (a) to what extent are pragmatic awareness and production teachable skills? (Kasper & Roever, 2005; Ohta, 2001); (b) how do (explicit and implicit) instruction and exposure to the L2 alone (e.g., in naturalistic language learning) compare in terms of effectiveness in facilitating the acquisition of L2 pragmatic norms? (Liddicoat & Crozet, 2001; Rose, 2005); and (c) do different instructional
approaches (e.g., implicit vs. explicit instruction) affect the acquisition of L2 pragmatic competence differentially? (Rose & Ng, 2001; Tateyama, 2001).

In addressing the key questions posed above, researchers in the area of developing L2 pragmatics have focused on many different aspects of the acquisition process. Several productive studies have been carried out in the context of child immersion education programs. Lyster (1994), for example, investigated the effects of a particular teaching method (functional-analytic teaching) on the sociolinguistic competence of French Grade 8 immersion students. In other studies, the focus has been on the discourse strategies that learners use in developing L2 pragmatic competence in the context of immersion instruction. For example, Kanagy (1999) investigated the acquisition of interactional classroom routines among English-speaking kindergarten-level immersion students learning Japanese.

Studies such as Kanagy’s (1999) and Lyster’s (1994) shed light on the general effects of formal instruction (immersion in particular) on child L2 pragmatic competence development, and have the added benefit of being analogous in terms of participants and the context of instruction to Swain (e.g., 1998) and Swain and Lapkin’s (e.g., 1995) research into the role of Output in the SLA processes of French immersion students. The present study is essentially a simultaneous examination of instructional effects on L2 pragmatic acquisition and the effects of Output on the acquisitional process, albeit with several key differences from Lyster, Kanagy, and Swain’s research efforts (e.g., a focus on adult ESL learners).

The Focus of the Present Study

In order to investigate the role of Output and classroom-based instruction in the development of L2 pragmatic competence, the present effort looks at the effectiveness of a particular method of enhancing interlanguage pragmatic development—video-based pragmatic instructional units along with output-focused activity—for improving pragmatic awareness and pragmatically appropriate production (Bou-Franch & Garcés-Conejos, 2003; Mir, 2001). Additionally, in response to Bardovi-Harlig and Hartford (2005) and Cohen (2004), who have called for a focus on more authentic data, the study includes not only a written discourse completion task (DCT), but also a somewhat more spontaneous oral DCT to enhance the reliability of the design. The inclusion of both written and oral production data may have the
further advantage of facilitating a fuller understanding of learner pragmatic development (Kasper & Roever, 2005; Martinez-Flor & Fukuya, 2005; Yuan, 2001). Finally, the inclusion of a receptive task, a pragmatic acceptability judgment task (PAJT; Garcia, 2004) allows for better triangulation of findings regarding the participants’ pragmatic development.

The primary issue addressed in this research effort is the efficacy of a specific type of output-focused instruction on the acquisition of developing L2 pragmatic competence. The larger question raised by the more specific research questions is whether Output functions in the case of pragmatics much as it has been shown to for the acquisition of morphosyntax. It is perhaps reasonable to expect that Output might have some effect on the acquisition of pragmalinguistic and sociopragmatic proficiency, but the question remains as to why this should be the case. What is needed is a coherent theory of developmental (or interlanguage) pragmatic acquisition that recognizes the complex nature of the process and accounts for its key components.

Among the elements of the L2 pragmatic competence acquisition process that need to be accounted for are the pragmalinguistic input available to learners, the developing pragmatic system that processes the pragmalinguistic forms that are taken in, and the output of pragmalinguistic forms that is produced by the learner. In this understanding of L2 pragmatic processing, sociopragmatic proficiency must be developed simultaneously to pragmalinguistic proficiency in order for there to be a balance of accuracy in form production and contextual awareness. In fact, a lack of sociopragmatic development has been posited as a possible problematic result of some efforts to teach L2 pragmatics explicitly (Kasper & Rose, 2002; Liddicoat & Crozet, 2001). Taking for granted that sociopragmatic development needs to take place, the focus of the present study remains overall L2 pragmatic competence, as indicated by the ability of learners to perceive and produce pragmalinguistic forms appropriately.

A theory of developmental pragmatic competence requires a relatively clear conceptual model of the acquisition process. Kasper (2001) reported that developing pragmatics has been described as either (a) a component that interacts with a grammar component within a larger model of communicative competence, (b) a form of information processing that resembles the learning of grammar, (c) an emerging competence within a sociocultural framework, or (d) a competence that develops simultaneously with cultural understanding in recurring situated events within a language socialization framework (p. 502). For the present study, a model of L2 pragmatic competence acquisition is adopted that assumes the first description offered by Kasper
(the communicative competence description). The understanding of pragmatic competence as a component of communicative competence that interacts with a grammatical component is consistent with the research of Bos, Hollebrandse, and Sleeman (2004), who proposed a significant role for an additional “mapping component” by which L2 learners acquire the ability to assign one or more possible surface (grammatical) forms to the underlying pragmatic meanings intended by the speaker, as well as that of Kasper and Rose (2002), who described a “pragmatics-grammar nexus” (p. 190).

The pragmatic acquisition model assumed here builds on the basic integrated psycholinguistic model of SLA proposed by Gass (1988), supported by VanPatten and Cadierno (1993) and elaborated upon indirectly in Izumi (2003). The model is realized within the framework of Swain’s (1985, 1995, 1998) Output Hypothesis, thus there is a strong role proposed for Output in the acquisition process. Figure 1 depicts the relationships among various aspects of the process of L2 (interlanguage) pragmatic competence acquisition.

![Figure 1. Model of Developing Interlanguage Pragmatic Competence Acquisition.](image-url)
The proposed model depicted in Figure 1 assumes that the acquisition of pragmatic competence in L2 is an iterative process (i.e., the learner’s pragmalinguistic output feeds back into the input that is then available to the learner’s developing IL pragmatic system). In the model, output of pragmalinguistic forms occupies an important place in the mechanism of L2 pragmatic competence acquisition. First, Output is proposed to contribute to the developing system of the learner as hypotheses about pragmalinguistic production possibilities are tested and either found to be accurate or wanting. Also, as learners compare their own output to acceptable pragmalinguistic forms in the input, the noticing function of Output is activated. In general SLA, it is theorized that in noticing the forms of their own utterances as well as those in the input, learners are theoretically being moved from semantic processing of language to more syntactic processing (Swain, 1993; Swain & Lapkin, 1995). For the purposes of L2 pragmatic development, this may mean that Output allows learners to move from focusing on outward pragmalinguistic forms associated with certain speech acts (e.g., recognizing that “thank you” may be used in response to a compliment) to a deeper processing of input (e.g., understanding that there are different ways to respond to compliments in different social and cultural situations). In addition, there is an interaction between Output and metapragmatic awareness indicated in the model. Finally, the output produced by learners serves to elicit and, to some degree, to shape the input that is produced by an interlocutor (indicated in Figure 1 by the broken line).

The grammatical component, indicated in the model as a “Grammar Halo” surrounding the learner’s developing system, corresponds to the grammar module proposed by Bos, Hollebrandse, and Sleeman (2004) and alluded to by Kasper (2001). The role for the grammatical component in interacting with the developing pragmatic component is two-fold, and includes the grammatical analysis of processed input as well as the consideration of candidate utterances for output of pragmalinguistic forms. This dual role of the grammar component of the learner’s developing system is depicted in Figure 1 by the encircling of the developing system by the aforementioned “Grammar Halo,” which serves both to filter input (what becomes available to the developing system as intake) and Output (what the learner is able to produce). Stainton (2005) posited that the primary distinction between grammatical (syntactic) processing and pragmatic processing is that syntactic control involves being able to link words (e.g., pronouns) to other linguistic referents, whereas pragmatic control involves the linking of a linguistic
expression with some non-linguistic concept, understood through the context in which the utterance occurs. In the sense in which Stainton distinguished syntactic and pragmatic processing, the primary difference seems to be in the type of anaphora that occurs with each. The ways in which a grammatical component might interact with the pragmatic component in the learner’s developing interlanguage system may be a matter of the grammatical component informing the pragmatic component as to the need for intra-sentential anaphora and the developing pragmatic component informing the grammar component as to items in the linguistic data that need to refer to extralinguistic, context-dependent elements. In any case, the important interaction between the developing pragmatic component and the grammar component in the model depicted in Figure 1 is consistent with the conceptualization of developing pragmatic competence as a component within a larger model of communicative competence (Kasper, 2001).

The Theoretical Significance of the Present Study

The assumptions upon which the proposed model of developing pragmatic competence acquisition is constructed must be grounded in careful research of the effects of not only input and learner variability, but also of Output on the demonstrable acquisition of pragmalinguistic forms. Among the assumptions underlying the model are that the process of L2 pragmatic competence acquisition is analogous to the general model of SLA accepted by many in the field of applied linguistics and SLA research (e.g., Gass, 1988) and that there may be an important role for Output in the acquisition of that competence. A review of the L2 pragmatics literature reveals that in most of the empirical studies that have been conducted, the focus of research has been confined to the roles of input generally, instruction (implicit and explicit) more specifically, and individual differences in the acquisition process. Little attention has been given to Output specifically in the context of the L2 pragmatic acquisition process. The dearth of research into the effects of Output (in production-oriented instruction) on developing interlanguage pragmatics thus motivates the present investigation.

The purpose of the present study to investigate whether Output in the context of instruction has a significant effect on developing L2 pragmatic competence has implications for SLA theory generally, and for the understanding of developing L2 pragmatics in particular.
VanPatten (2004) has claimed that evidence supporting a strong role for Output in SLA is insufficient, but Izumi (2003), Swain (2005) and others have continued to offer evidence that Output allows learners to achieve a level of linguistic awareness and development that they appear to be unable to attain without sufficient opportunities for Output. In interlanguage pragmatics research, Bardovi-Harlig (2001) has provided evidence that even given great amounts of time in the environment of the target language and culture, learners often fail to achieve target language-like pragmatic production. If a beneficial effect for Output in the context of instruction on the acquisition of L2 pragmatic competence is supported, the close relationship between pragmatic competence and general linguistic competence would suggest that Output is likely also an important part of SLA.

An additional potential theoretical contribution of the study is in the area of interlanguage pragmatic assessment and research design. The inclusion of multiple measures of developing L2 pragmatic competence (oral and written DCTs and a pragmatic acceptability judgment task, or PAJT) addresses Roever’s (2004) call for more research into the comparability of assessment instruments in L2 pragmatics research. In addition, because the performance and perception of speech acts are the outcome measures in the present study, there is a potential for the study to add to the body of speech act assessment research, discussed in depth in Cohen’s (2004) very helpful review. Finally, the inclusion of an analysis of the rubric used to rate the pragmatic acceptability of the study participants’ responses may further strengthen the validity and reliability of data collection methods in acquisition-focused L2 pragmatics research, in response to Boxer (2004) and others. The potential contributions to the understanding of L2 pragmatic acquisition assessment and research design add to the theoretical significance of the present study.

**Research Questions**

In light of the theoretical focus of the present study, the following research questions are posed:

1. Is there an effect for Output in the context of video-based pragmatic instruction (including output-focused tasks) on the developing L2 pragmatic competence of adult
ESL learners, as evidenced by the acquisition of pragmalinguistic forms (pragmatic speech acts) in appropriate contexts? Specifically, three sub-questions are posed:

1a. Is there a significant increase (pretest-posttest) for the PAJT measure?
1b. Is there a significant increase (pretest-posttest) for the written DCT measure?
1c. Is there a significant increase (pretest-posttest) for the oral DCT measure?

2. Is there a differential effect of Output on the two different pragmatic speech acts (directives and expressives)?
   2a. Is there a differential effect for Output on the perception of directives and expressives on the PAJT pretest and posttest?
   2b. Is there a differential effect for Output on the production of directives and expressives on the WDCT pretest and posttest?
   2c. Is there a differential effect for Output on the production of directives and expressives on the ODCT pretest and posttest?

3. Is the proposed rubric a reliable, valid measure for assessing the pragmatic acceptability of adult ESL learners’ linguistic production?

The first research question (including the three sub-questions) and the second question are explored in the primary data collection and analysis effort in the present study. The third question is dealt with in a separate analysis of the raters’ scoring of learners’ responses to the written and oral discourse completion tests.

**Summary**

The present study addresses a practical and theoretical need evident in the existing L2 pragmatics literature for research that continues to place the focus of developing L2 pragmatics research on the actual process of acquisition (Bardovi-Harlig, 1999; Kasper & Rose, 2002). Specifically, the study seeks to clarify the possible role of Output in the development of L2 pragmatic competence by adopting a proposed model of developing interlanguage pragmatic acquisition constructed within the framework of Swain’s (1985, 1995, 1998) Output Hypothesis. The theoretical significance of the present study consists primarily in its potential contributions to developing L2 pragmatics research, and SLA research in general, in two main areas: the proposed role of Output in instruction and the understanding of how speech acts are assessed.
Outline of the Dissertation

Chapter 1 has described the theoretical basis for the central assertion of the present study that instruction, particularly output-focused instruction, may be effective for the acquisition of pragmatic competence in a second language. Some initial evidence in the literature has been offered in support of the claim as well. The first chapter ends with the research questions addressed in the study and a glossary of relevant terms. Chapter 2 offers a more thorough review of the relevant L2 developing pragmatic and second language acquisition literature in motivating the variables of the study. Areas of focus in the second chapter include the notions of instruction, Output, and developing L2 pragmatic competence, as well as analysis of the reliability and validity of rating scales such as the DCT rating rubric proposed in the present study. The third chapter details the design and methodology of the research study. A review of the pilot study that preceded the present research effort is also included. In Chapter 4, the results of the quasi-experimental study and the analysis of the rubric used to assess the pragmatic acceptability of the participants’ responses are presented. Finally, the fifth chapter offers an analysis of the results and a discussion of the limitations of the study, as well as suggestions for future research in the field of developing L2 pragmatics research.

Definition of Terms

**Acquisition**: The gradual, often non-linear, internalization of a given linguistic form, as demonstrated by the ability of a speaker/listener to perceive and produce the form, which she was not previously able to perceive or produce (Doughty, 2003; Norris & Ortega, 2003).

**Communicative competence**: The overall ability of a speaker/listener to process and produce language and paralanguage effectively for the purpose of conveying information in real-world contexts (Cazden, 2001).

**Developmental pragmatics**: The study of learners’ developing interlanguage pragmatic competence, by means of which they are able to process and communicate desired meanings to other speakers of the L2 (Ohta, 2005).

**Discourse completion task** (DCT): A type of written or oral questionnaire that offers speakers the
opportunity to respond to a simulated situation that targets a specific speech act or set of speech acts (Yuan, 2001).

**Face:** The self-image that a person has and expects others to respect in public interaction (Yule, 1996).

**Grammar component:** A module within an overall model of communicative competence (labeled the “Grammar Halo” in Figure 1) that filters the input and output that a speaker/listener can process and produce (Kasper, 2001).

**Illocutionary force:** The aspect of a speech act that has to do with the speaker’s point, or intention, in producing the utterance (Yule, 1996).

**Input:** Linguistic stimuli (verbal or written), produced with the intention of communication, to which a speaker/listener is exposed (VanPatten, 2004).

**Instruction:** Formal methods designed to assist learners in the acquisition of a second language (Doughty, 2003).

**Intake:** A subset of input that is actually attended to by a speaker/listener and, in the case of a language learner, that is processed and comprehended by the developing system (VanPatten, 2002).

**Integrative processing:** the assimilation and organization of related forms into a coherent set, involving multiple operations performed together (Izumi, 2002; Robinson, 1995).

**Interlanguage pragmatics:** In a model of developing L2 pragmatic competence, the interlanguage pragmatic component is the locus of processing of pragmalinguistic forms, taking into account the sociopragmatic context (Bardovi-Harlig & Hartford, 2005).

**Metapragmatic awareness:** Explicit knowledge that a speaker/listener possesses about the forms and functions of pragmatic speech acts (House, 1996).

**Output:** Linguistic production, either oral or written, that is communicative in purpose (Swain, 1985).

**Pragmalinguistic forms:** Linguistic forms that are used to achieve certain speaker purposes (e.g., apologizing, complimenting, responding to requests, etc.) within the context of communication (Takahashi, 2005a).

**Pragmalinguistic proficiency:** The degree to which one is able to use appropriate linguistic forms to realize speech acts and their associated strategies (Leech, 1983; Thomas, 1983).

**Pragmatic competence:** The ability of a speaker/listener to achieve desired effects or convey
intended meanings in the course of communicative interaction (Kasper & Roever, 2005; LoCastro, 2001).

**Pragmatic filter:** A cognitive device that limits or constrains the pragmalinguistic intake entering a learner’s developing pragmatic system, theoretically in response, in part, to the sociopragmatic context in which the input was produced (by analogy with Krashen’s, 1985, affective filter).

**Pragmatic speech act:** An illocutionary act; i.e., a speech act considered specifically in terms of its connection to the communicative intent of the speaker who produced it (Kasper & Rose, 2002).

**Pragmatics:** The way in which speakers use language to achieve desired effects, make reference to implicit concepts, or convey intended meanings in communicative interaction in social and cultural contexts (Cummings, 2005).

**Second language acquisition (SLA):** The process by which a person learns to process and produce a language that is not her first language (Norris & Ortega, 2003).

**Socioculturally conditioned cues:** Linguistic and non-linguistic signals that allow a speaker/listener to know how to modify her own linguistic production in order to be pragmatically and sociolinguistically appropriate (Al-Issa, 2003; Ohta, 2005).

**Sociolinguistic competence:** The ability of a speaker/listener to produce and perceive linguistic forms in ways that are appropriate for the social situation in which the forms are produced (Canale & Swain, 1980; Cohen, 1996).

**Sociopragmatic proficiency:** The degree to which one is able to negotiate social situations involving social variables in pragmatically appropriate ways, including the ability to perceive and produce language that is sensitive to the context (Leech, 1983; Thomas, 1983).

**Speech act:** A communicative action, realized by means of spoken or written language and drawing on the social and cultural context, that reflects the intended meaning of the speaker or the speaker’s desired effect (Brown & Levinson, 1987; Searle, 1979).
CHAPTER 2

LITERATURE REVIEW AND MOTIVATION OF THE VARIABLES

The goal of the second chapter is to present a review of the relevant literature with regard to the central constructs and variables of interest in the present research effort, thus motivating the variables of the study. The treatment in this study is output-focused instruction in the context of a video-based instructional unit. The literature review provides an understanding of the nature of the variables of Output, pragmatic speech act type, and the acquisition of L2 pragmatic competence, as operationalized by written and oral pragmatic production and pragmatic perception. The discussion of previous research in the areas of second language pragmatic acquisition and Output indicate the need for the present study to explore the potential interrelationships among video-based pragmatic instruction, opportunities for output, and second language pragmatic acquisition.

Chapter 2 begins with a discussion and review of the literature concerning several constructs that are of central importance in the present study. First, the notion of pragmatic competence as it is understood in the context of this study is defined through an exploration of the successively more specific constructs of communicative competence, sociolinguistic competence, and pragmatic competence, followed by an examination of the interface of linguistic and pragmatic competence. The second chapter continues with a closer look at the issues around which the research questions of the study center.

Because the present study is fundamentally a test of the efficacy of a type of instruction on acquisition, the chapter next moves to a review of studies that have addressed the issue of how input and instruction affect SLA generally and developing L2 pragmatics more specifically. A discussion of research that has implications for the role of Output in SLA and L2 pragmatic development and acquisition follows. In addition, several studies that have implications for the assessment of developing L2 pragmatic competence are reviewed, in light of the third research question. The chapter closes with a discussion of research that treats the development of L2 pragmatic speech acts.
Defining Pragmatic Competence: Key Constructs

Communicative Competence

Swain (1985) explored communicative competence, considering it to be a construct that comprises three distinct components: (a) grammatical competence, (b) discourse competence, and (c) sociolinguistic competence. Ellis (1992) reported on the widely accepted idea that the development of communicative skills in the classroom facilitates SLA, finding that the L2 classroom setting might not always afford learners adequate opportunity to develop in all areas entailed in the concept of communicative competence. Deficiencies in one or more areas of communicative competence development in L2 such as those identified by Ellis may be attributable to a lack of opportunity for the learner to produce pragmatic speech acts in context, which may provide a second language learner with the ability to move beyond surface-level knowledge (e.g., formulaic expressions in response to some social situations) to a deeper level of analysis of pragmalinguistic forms (Martínez-Flor & Fukuya, 2005; Swain & Lapkin, 1995).

From a classroom discourse perspective, communicative competence may be viewed as a necessary element of the learner’s contribution to talk in the classroom setting. Cazden (2001) proposed communicative competence as that which a learner must develop in order to participate in the propositional, social, and expressive aspects of classroom language. Cazden also noted a connection between learners’ ability to self-correct and developing communicative competence, proposing that self-corrections are indicative of development on both the syntactic and the social plane as learners are better able to not only perceive inaccurate linguistic forms in their own output, but also to adapt their production to accommodate the needs of their interlocutors (peers or teachers).

Lazaraton (2004) studied the L2 communicative competence of a non-native speaking teacher of English as a Second Language (ESL) at an intensive English program (IEP) in the U.S. using a qualitative, conversation analytic (CA) approach. Using the inductive methods of CA, she identified one surface (grammatical) error made by the teacher in one hour of transcribed data, but noted that neither the teacher nor the students (English language learners) in the class noted the error or otherwise demonstrated any orientation toward the teacher’s non-nativeness. In that sense, Lazaraton argued, the findings reveal both something of the teacher’s
L2 communicative competence and the relationship between that competence and her performance in the classroom context as an instructor.

For the purposes of the present study, Long’s (1990) understanding of communicative competence as an aspect of overall language ability, along with linguistic competence, is useful. Long put forth the position that, particularly in the area of L2 learning and instruction, learners’ ability to understand what they are saying and why they are saying it in a given context is as important as achieving grammatical accuracy. Long’s concept of communicative competence is different, though not entirely inconsistent with, that of Celce-Murcia and Olshtain (2005), who, building on the earlier work of Canale and Swain (1980), Gumperz (1972) and others, identified the key aspects of communicative competence as linguistic, sociocultural, discourse, and strategic competence. Although Long’s idea of communicative competence is perhaps more narrow than some others (considering it to be complementary to, rather than comprising, linguistic competence), the important idea in both Long’s and Celce-Murcia and Olshtain’s conceptualizations of communicative competence is that it is a complex construct that encompasses developing sub-competencies. In an effort to move closer to a fitting and informed definition of pragmatic competence, a closer look at sociolinguistic competence, which has already been identified as a key component of communicative competence (e.g., Swain, 1985) follows.

**Sociolinguistic Competence**

Sociolinguistic competence as an aspect of communicative competence has received some attention in the SLA literature. In order to understand how the construct has been understood in the SLA literature, a few roughly comparable studies involving learners in a French immersion context are reviewed here. First, Swain and Lapkin (1990) analyzed two aspects of sociolinguistic performance: the appropriateness of an opening phrase in a note and the use of the conditional, also in the context of a written note. In addition, Swain and Lapkin described the use of subject pronouns by the participants (early and late French immersion students in Grade 10) in a job interview, and the production of requests, complaints and suggestions in the context of oral presentations. By scoring the students’ responses based on the presence or absence of formal features, the researchers sought to establish the learners’ ability to adjust their language use according to varying levels of formality.
The conclusion of the authors was that early immersion students outperformed later immersion learners in sociolinguistic speaking performance. Swain and Lapkin (1990) also noted that the early immersion students appeared to learn inductively from their teachers’ performance, whereas the later immersion students appear to have been relying on deductive means (i.e., attempting to follow explicit instructions given by the teacher). Swain and Lapkin’s findings suggest that sociolinguistic competence is fostered by authentic communication, and that length of exposure to the target language (or possibly age of first exposure) may have an effect on the ability of learners to internalize sociolinguistic forms. What these findings tell us about the nature of sociolinguistic competence in general is that classroom instruction alone may be insufficient to develop it (see also Canale & Swain, 1980, for further discussion of the place of sociolinguistic competence within the construct of communicative competence).

Swain and Lapkin’s (1990) relatively early study has been followed by many more recent works in the area of sociolinguistic competence. Dewaele (2004) investigated the sociolinguistic competence of 125 child French immersion students (L1 English) using a self-reporting questionnaire and an analysis of conversations between native speakers and non-native speakers of French (NS-NNS conversations). It was Dewaele’s concern to identify “sociobiographical variables” (p. 383), as well as situational factors, that might influence the use of pronouns of address (a sociolinguistic concern). Dewaele found that among the learners he tested, the use of pronouns was most affected by the status of the interlocutor (a social variable), more than by the age or gender of the speaker. One of the more surprising findings Dewaele had was that his learners tended to overuse the formal second person pronoun \( \text{vous} \), whereas learners in previous studies of French learners’ use of pronouns (e.g., Lyster & Rebuffot, 2002) had tended to overuse the informal address form \( \text{tu} \). Dewaele suggested that the difference could be accounted for by the fact that the learners in the other studies were younger than those in his study, and that the learners in those studies were using a communicative learning approach, involving large numbers of conversations with peers (requiring the use of \text{tu} ), and including a textbook that frequently featured informal conversations as examples. Those factors certainly could contribute to the sociolinguistic competence of learners.

Lyster (1994) targeted sociolinguistic competence directly as an outcome variable in a test of a functional-analytic teaching approach that involved learners’ practicing “sociostylistic variation” (p. 263). In a series of experiments involving a total of 106 Grade 8 French immersion
students, Lyster found that the functional-analytic approach utilized in the study did increase the sociolinguistic competence of his learners in at least three areas: (a) the ability to use the formal vous appropriately in various speaking situations, (b) the ability to use vous appropriately in formal letters, and (c) the ability to recognize appropriate French language use in context.

Sociolinguistic competence, as it has come to be understood in the SLA literature and specifically in studies of French immersion learners, entails both the ability to produce and perceive language accurately and to demonstrate a reasonable level of understanding of the social situations in which language occurs. It is just this intersection of linguistic and social factors, along with considerations such as politeness and reference, which leads to a discussion of pragmatic competence.

**Pragmatic Competence**

Watts (2003) proposed that pragmatic competence includes both the conversational maxims proposed by Grice (e.g., 1975; see Table 1) and rules of politeness, so that attempts to understand pragmatic competence should consider both speakers and listeners in conversational interaction in order to obtain an accurate view of the tension that is missing when researchers focus only on either speaker or listener alone. Watts’ politeness orientation toward understanding pragmatic competence is consistent with Bardovi-Harlig and Hartford’s (1991) status congruence approach to understanding learner pragmatics, which also drew on Grice’s conversational maxims in studying the speech acts produced by 21 non-native speakers (NNS) of English in interaction with native-speaking (NS) advisors or professors during academic advising sessions (authentic data). The study focused on examples of advice rejection produced by the learners, and the findings were that the NNS students largely followed native speaking students in accompanying instances of rejection with explanation, but often were not successful in maintaining the balance of status with their NS interlocutors because they chose inappropriate status-preserving strategies when rejecting the advice. For example, one Korean speaker of English as a second language rejected her graduate school advisor’s advice to take two different courses by using apparently contradictory reasons: “I checked the textbook but—all the information is just new to me” and, for the second class, “But I checked the textbook and it’s exactly the same” (p. 54). The apparently contradictory explanations, possibly intended by the Korean learner to avoid directly refusing the advice, instead resulted in confusion for the advisor and, eventually, dismissal of the learner’s objections by the advisor.
The failure to adhere to conversational maxims, though not always indicative of a lack of pragmatic development in the L2, is a consistent problem for L2 learners. Certainly, familiarity with the ways in which the conversational maxims, as proposed by Grice (1975) are commonly realized in American English would be advantageous to ESL learners who will be in regular social and academic interaction with native speakers. Grice’s maxims are outlined in Table 1.

<table>
<thead>
<tr>
<th>Maxim</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>Interlocutors should be as informative as necessary in their contributions,</td>
</tr>
<tr>
<td></td>
<td>but not more than is necessary.</td>
</tr>
<tr>
<td>Quality</td>
<td>Interlocutors should say things that they believe to be true, or that</td>
</tr>
<tr>
<td></td>
<td>evidence suggests is true.</td>
</tr>
<tr>
<td>Relevance</td>
<td>The contributions of the interlocutors should be relevant to the direction</td>
</tr>
<tr>
<td>(Relation)</td>
<td>and purposes of the conversation.</td>
</tr>
<tr>
<td>Manner</td>
<td>Interlocutors should make their contributions in an orderly way, in keeping</td>
</tr>
<tr>
<td></td>
<td>with expected conversational norms in the culture.</td>
</tr>
</tbody>
</table>

In discussing conversational maxims, it is important to note that there is no expectation that speaker/listeners will always follow the maxims. Rather, the expectation is that interlocutors will adhere to a basic “cooperative principle” (Grice, 1975). This means that when a maxim is not followed, there is some implicature (implied meaning) being conveyed, and the interlocutors would be expected to negotiate through talk to understand what the implicature is, and to maintain a kind of conversational balance, thus following the cooperative principle.

Bardovi-Harlig and Hartford’s (1991) findings in their study of advising sessions are similar to those of Davies and Tyler (2005). In their examination of the pragmatic discourse strategies used by a Korean international teaching assistant (ITA) at an American university, Davies and Tyler concluded that the difficulties faced by the Korean ITA they studied were
related to the inability of the learner, who was teaching native English speaking university students in a physics lab, to employ appropriate pragmatic discourse strategies. Through analysis of naturally occurring speech, the researchers found that the ITA was utilizing discourse strategies that did not fit with expected American English pragmatic norms. The conclusion reached by Davies and Tyler was that the ITA tended to choose unproductive discourse strategies (e.g., adopting an assertive/deductive stance in a confrontation with an undergraduate student in his class) in response to stressful situations in the unfamiliar environment of the U.S. university classroom.

The unsuccessful approach taken by the ITA in Davies and Tyler’s (2005) study contrasts with what Bardovi-Harlig and Salsbury (2004) found in a longitudinal study of the disagreements produced by ESL learners in conversations with NS of English over a one-year period. In that study, the learners developed the ability, over time, to disagree while employing such status-preserving strategies as increasing the amount of talk (elaboration), including some agreement in the oppositional talk, and manipulating turn-taking structure to try to avoid direct conflict. The cross-section of language, context, social norms and expectations that makes up pragmatic competence is evident in these studies.

**Speech acts.** A central concept in the study of general pragmatic competence is the speech act. Boxer (2002) discussed speech acts as a micro phenomenon within sociolinguistic analysis. Kasper (2004b) investigated the performance of speech acts in the context of authentic language use by using the analytical techniques of conversation analysis (CA). In the study, Kasper examined several extracts of authentic language to identify the nature and purposes of repeated questions as used by NNS participants. Kasper noted that speech act research typically relies on the analysis of social variables such as relative power, social distance, etc. to understand the formatting of given speech acts (Brown & Levinson, 1987). In this study, however, Kasper concluded that understanding the makeup of repeated questions (directive speech acts), required combining the usual speech act research methodology with the closer text-focused analysis of the CA approach.

Speech acts have been categorized and classified in different ways over time, from Austin (1962) to Leech (1983) and more, but a commonly accepted classification of speech acts (or illocutionary acts) was provided by Searle (1979). Searle’s classification of speech acts, as outlined and presented in Mey (2001, pp. 120-123), includes categories that are oriented toward
the pragmatic intentions of speakers. In that sense, Searle’s classification system may be regarded as primarily concerned with the illocutionary force of speakers’ speech acts (i.e., what speakers intend to accomplish by producing given speech acts). Searle’s categories, along with explanations for each, are detailed in Table 2.

<table>
<thead>
<tr>
<th>Category</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representatives</td>
<td>assertions about the world based on the belief of the speaker (e.g., “Today is Monday.”)</td>
</tr>
<tr>
<td>Directives</td>
<td>attempts to get a listener to do something in response to the utterance (e.g., advice, invitation, or a request such as: “Could you pass the salt?”)</td>
</tr>
<tr>
<td>Commissives</td>
<td>obligations that a speaker places on herself (e.g., a promise)</td>
</tr>
<tr>
<td>Expressives</td>
<td>expressions of the inner thoughts or feelings of the speaker (e.g., an apology, request refusal, or response to a compliment)</td>
</tr>
<tr>
<td>Declarations</td>
<td>formal statements by which a speaker declares something to be true or accepted (e.g., “I pronounce you husband and wife.”)</td>
</tr>
</tbody>
</table>

The Linguistic-Pragmatic Interface

Bachman (1990) proposed pragmatic competence as a component of a larger concept of language competence, along with organizational competence. Bachman’s conception of pragmatic competence within a communicative language ability framework is consistent with that of Roever (2004), who discussed pragmatic competence as entailing both conventional formal means of expressing intentions (pragmalinguistic knowledge) and social norms and
obligations (sociopragmatic knowledge). In Roever’s description, there are both situation-dependent (contextual) and non-contextual aspects of pragmatic competence. In that sense, it may be stated that linguistics and pragmatics interface at the place where formal linguistic means are required to realize the pragmatic intentions of a speaker (and the subsequent perception and interpretation of the act). Douglas (2004) discussed discourse domains as a speaker/listener’s responses to the situational (i.e., social and cultural) and linguistic environment. Pragmatics and linguistics again interface in this conceptualization of interaction between the individual and the context.

Perhaps the clearest example of the interface of linguistics and pragmatics is the concept of pragmalinguistic proficiency. An example of the operationalization of this construct is Takahashi’s (2005a) study of pragmalinguistic awareness and its possible relationship to motivation and proficiency. In that study, Takahashi conceived of pragmalinguistic awareness as possibly being analogous to awareness or noticing in the general language competence. In the course of the study, it becomes clear that pragmalinguistic awareness, for Takahashi, has to do with the noticing of specific forms that have pragmatic meaning. In Takahashi’s study, then, the intersection of surface form and the attaching context-dependent intention or meaning again emerged as a productive place for the exploration of pragmatic competence. It is because an understanding of form and intention in context are so integral to pragmatic competence that developing pragmalinguistic proficiency is such a useful gauge of development in a learner’s interlanguage pragmatic system.

Facilitating L2 Pragmatic Competence: Key Questions

This section of the second chapter reviews studies that have relevance for the key questions of interest in the present study. Assuming a model of developing L2 pragmatics as depicted in Figure 1 in the first chapter, this section features studies that can contribute to an understanding of the basis on which the research questions and hypotheses of the present study have been constructed. The section begins with a review of works that have focused on input and instruction as factors in the acquisition of L2 pragmatic competence.
The Effects of Instruction on Acquisition

Because the present research effort is fundamentally a test of the effects of instruction on acquisition, it is logical to move at this point into an exploration of the relevant literature addressing this phenomenon. Eslami-Rasekh (2005) has reported that evidence from ESL/EFL classrooms indicates that even English language learners (ELLs) at advanced levels of overall linguistic competence often fail to produce or interpret pragmatic intentions correctly. Eslami-Rasekh’s conclusion regarding ELLs echoes the earlier findings of Olshtain & Cohen (1990), who found that intervention was facilitative of acquisition of some pragmatic forms by learners who, despite overall advancement in linguistic proficiency, still exhibited problems with the targeted forms and strategies. There is thus some empirical basis, if not universal support, for the position that pedagogical intervention may facilitate some learners’ acquisition of at least some aspects of L2 pragmatic competence.

Instruction is conceptually a means of facilitating the assimilation of input. Thus, in the following discussion and review of studies related to instruction and developing L2 pragmatics, the role of input in developing pragmatic competence is explored first, followed by a discussion of instruction in SLA more generally and finally, a review of studies that have investigated instruction as it relates specifically to developmental pragmatics.

Input and Developing Pragmatic Competence

Before considering the impact of instruction on developing pragmatic competence in the existing literature, it is important to look first at the suggested roles for input in the pragmatic competence development process. The importance of input as a driver of second language acquisition more generally (VanPatten, 2004) and the prominent place of input in the proposed model of interlanguage pragmatic competence acquisition (Figure 1) make a discussion of input an important part of the upcoming discussion of Output in the context of instruction and its possible effects on developing L2 pragmatic competence. In short, without an appreciation of input in the acquisition process, it is not possible to understand the need for and potential benefits of instruction or Output.

Some input-focused studies of developing L2 pragmatics have explored the effects of environmental exposure over a period of time on the acquisition of pragmatic competence. Matsumura (2003), for example, looked at the relationship between amount of exposure to English and developing pragmatic competence in English of 187 Japanese ESL learners in an
exchange program in Canada. Matsumura’s finding that length of exposure only moderately correlated with learner performance on a posttest DCT (based on the speech act of advice giving) indicates that a large amount of input may not be sufficient in and of itself to explain the acquisition of productive pragmatic competence, just as Swain (1985) concluded that copious amounts of input alone could not ensure accuracy in morphosyntactic production, likely because of either insufficient quality (i.e., the input in classrooms contains a restricted range of forms and semantic domains) or because learners’ attention was not sufficiently drawn to salient features in the input.

Input made available through the environment does, however, appear to affect learners’ perceptions of pragmatic appropriateness. In a study contrasting ESL and EFL learners’ awareness and evaluations of pragmatic violations, Bardovi-Harlig and Dörnyei (1998) found that whereas 482 EFL learners in Hungary and Italy (and their 53 non-native English instructors) were less aware of pragmatic infelicities and tended to rank such violations as less important than grammatical errors, 173 ESL learners in the United States (average length of residence, about 5 months) tended to rank pragmatic violations as much more serious than grammatical miscues. Bardovi-Harlig and Dörnyei attributed the difference in awareness and judgment to the difference in residence between the two groups.

Environmental exposure may be a rather nebulous concept by which to judge input, but there have been acquisition-focused L2 pragmatics studies that have examined the effects of input in much more controlled settings. Takahashi (2001), for instance, looked at the extent to which input enhancement facilitated the acquisition of English request forms by 138 Japanese university students. The research questions were whether the degree of input enhancement would affect acquisition of the target request strategies, and whether the type of input condition would affect learners’ confidence in producing their own request forms (p. 174). A classic pretest-posttest quasi-experimental design was employed, and the discourse completion test (DCT) was the instrument utilized to elicit data for the pre- and posttests (p. 175). The four basic conditions of input that were established were explicit teaching (ET), form-comparison (FC), form-search (FS), and meaning-focused (MF).

Takahashi (2001) found that the ET condition enabled the learners to produce more target-like forms than the other conditions across all request situations. Also, learners in the ET and MF conditions exhibited a greater degree of increase in their confidence in formulating
request expressions (p. 189). In a qualitative analysis of the results, it was proposed that the apparent mastery of the ET group over the target request forms might not be stable or likely to continue, though a delayed posttest would be required to test this latter conclusion (p. 192).

Takahashi’s (2001) final conclusion was that L2 pragmatic competence cannot be enhanced with positive evidence through input enhancement alone, though other focus-on-form (FonF) techniques might be of use and should be explored in future research (p. 198). Although he found that the explicit enhancement of input facilitated acquisition of the relevant pragmalinguistic forms (based on DCT responses by the learners), it seems plausible that the effectiveness of the enhancement was due largely to the fact that the information provided targeted the learners’ metapragmatic awareness directly, and so it is difficult to conclude that the findings necessarily support input enhancement generally.

Taken together, the findings of research regarding input and developing L2 pragmatics appear to support the notion that the function of input here is analogous to its role in the more general process of SLA, just as Takahashi (2005a) concluded that pragmatic awareness is analogous to awareness in the more general language competence of a speaker. In fact, the correspondence between Matsumura’s (2003) results and those of Swain (e.g., 1985) suggests that there are empirically motivated grounds for investigating the possible effects of Output in the L2 pragmatics acquisition process as well.

**Instruction and SLA**

Doughty (2003) reported that the debate over whether and how instruction affects SLA eventually comes down to the two fundamental question of whether SLA involves primarily implicit or explicit processing and whether implicit or explicit instruction is more effective in facilitating the acquisition. Although the ability of learners to process language may be generally equated with acquisition, identifying evidence of acquisition is not a simple matter. A brief survey of studies that have investigated the effects of formal instruction on SLA reveals that instruction does appear to have an effect on SLA, although the exact nature of the effect may not be clear for all aspects of language. Norris and Ortega (2000) in a large meta-analysis of some 77 studies covering a period of 18 years (1980-1998) found that formal L2 instruction had larger effect sizes than simple exposure to L2 or non-focused treatments. Also, they found explicit instruction, defined as that which either incorporates explanation of rules (deductive or metalinguistic approach) or requires learners to focus or give attention to particular forms, to be
more effective than implicit instructional approaches, defined as instruction that does not involve rule explanation or instruction to attend to forms (p. 437). A thorough study of the effects of instruction in the general SLA literature is beyond the scope of the present study, but two particular areas of research have produced promising findings with some relevance to the present research effort, and are reviewed here.

First, in the area of focus-on-form based instruction (FonF), Doughty and Varela (1998) asked whether incidental, form-based instruction would help 34 ESL learners to improve in their ability to use past tense forms accurately in oral and written tasks. All the students (experimental and control groups) had the opportunity to complete oral and written science lab reports immediately after and then again about two months after the treatment period. Doughty and Varela analyzed the data, finding that the control group failed to make any maintained gains in target-like use of targeted past tense forms, whereas the experimental group, which had received incidental, FonF instruction (including, among other techniques, frequent use of recasts), made significant improvements that were maintained on the post-test two months after the treatment period. The positive findings concerning FonF instruction are consistent with Long’s (1991) position that focusing on forms within meaningful communicative acts is more effective than viewing forms as discrete units to be accumulated (out of context).

Another promising area of research in the area of instructional effects on SLA has been processing instruction. An exemplary study in this area of research is VanPatten and Oikkenon (1996). In this study of high school level learners of Spanish as a foreign language, the control group received explicit information and practice with structured input activities over a period of four days. Next, an “explicit information only” experimental group were given the explicit information, but not given any practice of any kind. Finally, a “structured input activity only” group received the input-based activities (again, over four days), but received no explicit information about pronouns, even as feedback for the activities. The third group, importantly, did no practice in output. VanPatten and Oikkenon found that the processing instruction (control) group and structured-input only groups increased in performance after treatment, but the explanation-only group did not. The findings support the efficacy of processing instruction for facilitating SLA, as demonstrated by learner gains on pre- to posttest interpretation and production tests.
Instruction and Developmental Pragmatics

Formal instruction has long been posited to have a beneficial effect on pragmatic development (e.g., Kasper & Rose, 2001; Olshhtain & Cohen, 1990). Exploring the effects of instruction in an authentic classroom-based study, Ohta (2001) looked at longitudinal development in the use of alignment expressions by American learners of Japanese as a foreign language (JFL). Her findings were that the learners’ development in the use of the target forms was related to the use of those pragmalinguistic forms by the teacher. Based on extensive field notes of the language of the classroom, Ohta was able to propose a six-stage developmental scale for the acquisition of the relevant speech acts, which could be useful for future research into the acquisition of pragmatic competence in the JFL context and possibly adapted to other L2 classroom contexts. Ohta’s finding that pragmatic usage by the learners in her study was essentially an emergent phenomenon, reflecting an apparently internal developmental schedule of pragmalinguistic forms, has implications for a possibly different effect of instruction on different categories of pragmatic speech acts.

Tateyama’s (2001) study of 27 learners of JFL yielded different results from those of Ohta (2001). In a study designed to contrast an explicit instructional approach to teaching Japanese pragmatics with a more implicit (input exposure only) approach, Tateyama found that learners in the two groups did not differ significantly in their performance on a multiple choice instrument designed to elicit pragmatic acceptability judgments. In four different sessions over eight weeks (average of one session every two weeks), the 27 participants were exposed to video clips that featured the targeted pragmatic speech acts, with the explicit instructional group also receiving teacher comments and handouts on how to use the pragmalinguistic forms and the implicit group receiving no additional support. Although no significant difference in performance was observed, one interesting result that emerged from the study is that on a secondary measure—an interactive role play—the learners who received explicit instruction actually performed worse than the implicit learners. Tateyama hypothesized that this unexpected finding was the result of “overpoliteness” on the part of the explicit learning group. This potentially negative effect of instruction is certainly one that must be guarded against, as pragmatic competence involves not only the avoidance of inappropriate speech acts, but also the avoidance of overapplication of politeness strategies. The cause of the lack of significance may
also have been that the instructional treatment, four sessions spanning eight weeks, was not sufficiently concentrated to maintain any effects for the explicit instruction.

Lest it be concluded that explicit pragmatic instruction is harmful to learners, the findings of Rose and Ng (2001) need to be considered. In a study of 48 university students in Hong Kong, the researchers found that explicit instruction in compliments and compliment responses in English enabled the majority of learners to demonstrate some gains on a written DCT instrument. In fact, Rose and Ng proposed that the implicit instruction group in their study appeared to have been harmed by their instruction sessions, presumably because the implicit nature of the teaching raised difficult sociopragmatic questions without offering corresponding pragmalinguistic responses.

Further evidence for the advantage of formal instruction in L2 pragmatic development has been provided by Bardovi-Harlig (2001). In her study of second language learners, Bardovi-Harlig obtained findings that strongly suggest that learners who do not receive pragmatic instruction fail to approximate native-speaker usage and comprehension of pragmatic forms. It is possible that the reason that learners who do not receive explicit instruction fail to approximate native-speaker pragmatics is that the metapragmatic knowledge that is needed to advance beyond a theoretical non-native ceiling simply cannot be developed through implicit exposure. An interesting question is whether it is the explicit instruction itself or the opportunity afforded to learners to push themselves in terms of hypothesis testing and the production of pragmalinguistic forms that provides the benefit.

It appears that in order for instruction to benefit learners in terms of pragmatic development, it should balance implicit and explicit presentation of information, and, most importantly, be embedded in context, as demonstrated by the success of Ohta’s (2001) classroom JFL learners. The importance of context is supported by the findings of Kim and Hall (2002), who investigated the EFL pragmatic development of four Korean children participating in an interactive reading program that involved a great deal of dialogic interaction with adult tutors. Over a period of several months, Kim and Hall found that the use of successful discourse management strategies and pragmatic formulas increased among all four children, despite a lack of any explicit instruction.

In a qualitative examination of the effects of instruction on the ability of 49 Japanese EFL learners to notice request forms in the target language, Takahashi (2005b) found that those
assigned to a form-comparison instructional group tended to outperform those assigned to a form-search instructional group, both during the treatment and on a posttest measure. During the four ninety-minute treatment sessions, Takahashi’s participants either filled out open-ended DCTs for comparison with English-speaking actors in role-play activities (the form-comparison group) or identified differences between native and nonnative speakers’ transcribed role plays (the form-search group). Because the form-comparison group members were engaged in comparing their own responses to a DCT with those forms produced by NS of English in a role play, whereas the form-search group were simply comparing the NS requests from the role play to those produced by other Japanese learners of English and transcribed for them, the two groups were engaging in fundamentally different types of activities. The distinction suggests the possibility (unexplored in Takahashi’s study) that Output played a role in the advantage enjoyed by the form-comparison group. Takahashi’s results further suggest that if some instructional treatments are more effective than others, the larger question of whether instruction generally has an effect on developing L2 pragmatics should be answered affirmatively.

The findings of the studies on instruction and developing second language pragmatics reviewed above suggest that the explicit/implicit dichotomy may not be as useful as a context-embedded/context-poor distinction when considering the effects of formal instruction on interlanguage pragmatic development (Tateyama, Kasper, Mui, Tay, & Thananart, 1997). This contention is supported in Brown and Levinson’s (1978) classic work on politeness, in which it was noted that children experiencing first language acquisition acquire politeness, particularly those aspects of politeness that do not involve complex linguistic forms, naturally from the context in which they are embedded. Also, Papafragou & Musolino (2003) noted that when children in an experimental study of pragmatic scalar implicatures (e.g., the ‘some’-‘all’ distinction) were aware of the purpose and context of the utterance, their performance was significantly better than when that context was lacking. In the present study, the context of pragmatic speech acts was provided for the participants explicitly on each of the test instruments and on the pragmatic video vignettes utilized during the instructional treatment as well. The extent to which the context-embeddedness of the items on the instruments and in the vignettes enabled the learners to make use of the output-oriented instruction was not explored directly in the present study, but was indirectly measured in the testing carried out. Table 3 summarizes the findings of the studies on instruction and pragmatics reviewed above.
Table 3  

*Studies on Instruction and Developing L2 Pragmatics*

<table>
<thead>
<tr>
<th>Study</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olshtain &amp; Cohen (1990)</td>
<td>French-speaking primary students in intensive ESL programs outperformed same-grade students in standard ESL programs in listening, reading, oral fluency, and positive attitudes toward English.</td>
</tr>
<tr>
<td>Tateyama, et al. (1997)</td>
<td>Beginning JFL learners receiving explicit instruction on the use of pragmatic routines (e.g., <em>sumimasen</em>) outperformed implicit learners on role-play and multiple choice tasks targeting formal expressions.</td>
</tr>
<tr>
<td>Ohta (2001)</td>
<td>American JFL learners’ use of alignment expressions developed in line with teacher’s use of the forms; results reflected a possible internal order of pragmatic development.</td>
</tr>
<tr>
<td>Rose &amp; Ng (2001)</td>
<td>University students in Hong Kong who received explicit instruction in compliments and responses outperformed implicit learners on a DCT; the implicit learners may have been harmed by their instruction.</td>
</tr>
<tr>
<td>Tateyama (2001)</td>
<td>Found no significant difference between learners who received explicit and implicit (input-only) instruction on a MC judgment test; implicit learners outperformed explicit learners on a role play task.</td>
</tr>
<tr>
<td>Takahashi (2005b)</td>
<td>Japanese EFL learners in a form-comparison instructional group outperformed those assigned to a form-search instructional group, both during the treatment and on a posttest measure.</td>
</tr>
</tbody>
</table>

**The Effects of Output on Acquisition**

The idea that Output promotes acquisition is, in some ways, intuitive, but is not without challenges with respect to SLA theory. VanPatten’s (2004) claim that successful SLA cannot be claimed to be dependent in any significant way on Output places the burden of proof on SLA researchers working within the framework of the Output Hypothesis to provide empirical evidence of a strong role for Output in the acquisitional process. Such evidence has been offered by researchers such as Izumi (2002, 2003) and Swain (1985, 1995), but differing interpretations of findings allow those who do not view Output as a strong facilitator of SLA (particularly in
comparison with input) to remain skeptical. A review of works that have focused on the role of Output in general SLA follows.

**Output in SLA**

Before discussing those studies that have explored the roles of Output in developing L2 pragmatic competence, a look at Output in the more general SLA process is beneficial, largely because the acquisition of L2 pragmatic competence may be analogous to the more general acquisition process. Arguments for this analogy include the proposed “pragmatics-grammar nexus” (Kasper & Rose, 2002, p. 190) mentioned earlier, as well as the connections between general SLA and the acquisition of L2 pragmatic competence suggested by efforts (e.g., Kasper & Schmidt, 1996) to move developing L2 pragmatics research to ask questions of relevance to general SLA theory and research.

Representative of output-focused research in SLA is Swain and Lapkin’s (1995) study of 18 Grade 8 French immersion students. In the study, the researchers sought to answer three basic questions: (a) Do learners notice their own gaps when producing L2? (b) If they are aware of the gaps, does this trigger cognitive processes that contribute to SLA? and (c) Do the learners engage in grammatical analysis of their output when trying to solve linguistic problems? In order to address these questions, Swain and Lapkin carried out a writing task with the learners. To do this, one of the researchers met with each learner and instructed them in the procedure for performing a “think-aloud” task (Ericsson & Simon, 1993) while writing an article about a common theme.

On analyzing the data that resulted from the think-alouds, Swain and Lapkin (1995) did identify evidence of learners’ noticing their gaps. Also, with regard to their second research question, the think-aloud revealed that the learners made use of such cognitive processes as extending their L1 knowledge to L2 interpretation and engaging in hypothesis testing. All of these things took place, it was noted, without any external feedback from the teacher or the researchers. The suggestion is clearly that Output has the potential to influence the general SLA process. In light of the focus of the present study on developing L2 pragmatics, the question naturally emerges as to whether Output plays a similar role in learners’ acquisition of L2 pragmatic competence. If developing second language pragmatic competence can be shown to be positively correlated with opportunities for output, a case may be made for the incorporation of
output-based instruction in the second language classroom. Before addressing that question, however, a closer examination of the Output Hypothesis is in order.

**The Output Hypothesis**

In the last ten to fifteen years, the construct of Output has been proposed as a possible contributing factor to second language acquisition because of its ability to promote noticing (Izumi, 2002), hypothesis testing (Mackey, 2002; Mackey, Gass, & McDonough, 2000), and metalinguistic reflection (Swain, 1995, 1998, 2005). Taken together, these three functions reflect the role of Output in actual language learning situations. The central claim of the Output Hypothesis, as articulated by Swain (2005), is that “the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning” (p. 471). The hypothesis, then, does not deny the importance of input, but reacts to some sweeping claims of the Input Hypothesis (Krashen, 1985) that argued for comprehensible input as the only necessary condition for SLA to occur.

Importantly, the Output Hypothesis argues for a fuller understanding of the SLA process, taking the theoretical burden for acquisition off of input alone and thus allowing a view of acquisition and the developmental process that acknowledges the potential advantage of approaches such as processing instruction (e.g., VanPatten, 2002) and focus on form (FonF; e.g., Doughty & Varela, 1998) for enhancing noticing while recognizing that the acquisition process is complex and that other phases of the process may be targeted for instruction profitably as well. According to the Output Hypothesis, the production of language by a learner is not merely the result of acquisition, but is rather a critical contributor to acquisition. In fact, research has supported the notion that when second language learners are exposed to large amounts of input but have inadequate opportunity to produce the target language, acquisition is short circuited to some extent in terms of morphosyntactic accuracy (e.g., the French immersion studies; Swain, 1985; Swain & Lapkin, 1986, 1995, 2002).

The locus of the breakdown in accuracy during the SLA process has been identified as learners’ inability to “notice the gap” when in the L2 environment (Schmidt & Frota, 1986). In the framework of the Output Hypothesis, noticing is one of the principal functions of Output. In a controlled experimental study of 61 adult ESL learners’ ability to notice English relative clauses, Izumi (2002) explored the potential effects of Output and input enhancement on noticing. Of the participants, 47 participated in the treatment, with the other 14 learners serving
as a control group and taking only the pre- and posttests. The treatment consisted of a text reconstruction activity in which the learners first read the text for understanding, then either reconstructed the text (for the Output groups) or answered non-comprehension oriented extension questions (for the nonoutput groups), and finally wrote a recall summary in L1 to demonstrate comprehension. Within the treatment, some learners received both opportunities to reconstruct the text and received enhanced input in their second exposure to the reading materials, whereas others received only enhanced input, but no opportunity to reconstruct (answering questions instead), so that Izumi could determine whether Output or enhanced input made a greater contribution to noticing.

The pretest and posttest measures used by Izumi consisted of four instruments designed to measure the learners’ understanding and ability to use English relative clauses: (a) a sentence combination test, (b) a picture-cued sentence completion test, (c) an interpretation test, and (d) a grammaticality judgment test. On analyzing all of the data (including the descriptive statistics for the learners’ note-taking and text reconstruction scores), Izumi concluded that the input enhancement did have an effect on noticing, but that the two Output groups showed significant improvement from pretest to posttest, whereas those with no opportunity to reconstruct the text did not show significant improvement. Thus, the only claims that could be made definitively based on the results were that input enhancement was better than no enhancement for promoting noticing, but that opportunity for output was the most important factor of all in predicting improvement from pretest to posttest.

The three functions proposed for Output in Swain (1995) are: (a) the noticing/triggering function, (b) the hypothesis testing function, and (c) the metalinguistic (reflective) function. These three functions work to enable learners to develop an awareness of their own developing systems at a level that mere exposure to input does not permit, moving the learners from language processing on a semantic level to processing on a more syntactic level (Swain & Lapkin, 1995). Importantly, it is also claimed that the functions of Output can operate even without explicit feedback from interlocutors. The question of whether these same functions are evident in the acquisition of L2 pragmatics is one that is explored further in the present study.

Output and Developmental Pragmatics

A valid question concerns the extent to which the Output Hypothesis, which appears to apply in the case of morphosyntactic acquisition (Izumi, 2002; Swain, 2005), is a valid construct
for application to instructed L2 pragmatic development. Kasper and Roever (2005) pointed out that the complexity of developing pragmatic competence means that exposure to input alone is insufficient to promote pragmatic development in a new language (p. 318). This observation is clearly supported by a consideration of the task before a language learner in terms of pragmatic development.

First, the learner must learn to master the linguistic system (morphosyntax) while simultaneously developing the ability to recognize the social and cultural implications of the forms that are being learned. Thomas (1983) usefully distinguished between pragmalinguistic failure and sociopragmatic failure when accounting for pragmatic errors, and Leech (1983), in the same year, applied the same terms to the two central aspects of overall pragmatic competence. The distinction has come to be broadly applied so that now pragmalinguistics and sociopragmatics may be regarded as distinct, though closely related and possibly overlapping, aspects of pragmatic competence (e.g., Rose, 2001; Takahashi, 2005a, who tested specifically for pragmalinguistic awareness).

Essentially, pragmalinguistic proficiency refers to a learner’s ability to use acceptable language to realize speech acts and their associated strategies. Sociopragmatic proficiency, on the other hand, refers to a learner’s ability to negotiate social situations involving social variables (e.g., age, gender, differences in status) in pragmatically appropriate ways (Bardovi-Harlig & Hartford, 1991; Kasper & Roever, 2005). Also, Rose (2000) used the term pragmalinguistic proficiency to refer to a learner’s ability to choose the appropriate speech act to fit the context and to make use of suitable linguistic forms to realize the speech act. He also referred to sociopragmatic proficiency as the ability to perform illocutionary acts according to social norms of acceptability and appropriateness.

In the present study, the focus of research is developing pragmalinguistic proficiency, or the emerging ability to make use of correct linguistic forms to enact appropriate speech acts in a language other than one’s first language. However, it is critical to note that the focus on pragmalinguistic forms does not ignore the importance of the sociopragmatic development that should be occurring along with the developing pragmalinguistic proficiency. Rather, the focus on pragmalinguistic proficiency within developing L2 pragmatic competence is a natural outgrowth of the conceptualization of the pragmatic processing model that views Output as primarily oriented toward pragmalinguistic forms, which also are the most observable evidence for the
development of L2 pragmatic competence. Furthermore, for the purposes of this study, Thomas’ (1983) distinction between pragmalinguistic and sociopragmatic phenomena is assumed to be comprehensive in that all pragmatic phenomena are held to be classifiable as either pragmalinguistic or sociopragmatic in nature and that all pragmatic phenomena are carried out within the larger construct of communicative competence, as distinguished from linguistic competence alone (Kasper, 2001; Long, 1990; Yule, 1996).

In terms of the task facing the second language learner, then, not only must the linguistic and pragmatic systems in and of themselves be mastered, but the interaction between the grammatical component and the pragmatic component within the learner’s overall developing communicative competence must function efficiently and with minimal interference from potential problematic factors such as lack of attentional resources or simple lack of metapragmatic awareness or noticing (Kasper & Schmidt, 1996). When one considers that SLA generally is a daunting task for learners, it comes as little surprise that the acquisition of L2 pragmatic competence, a parallel phenomenon, should be equally challenging. If there are parallels between SLA generally and the development of L2 pragmatic competence, there may well be a role for Output not only in SLA (i.e., in terms of morphosyntactic development), but also in connection with a learner’s developing L2 pragmatics. A summary of the potential effects of Output on developing L2 pragmatic competence in terms of Swain’s three proposed output functions for SLA, is detailed in Table 4.

<table>
<thead>
<tr>
<th>Output Function</th>
<th>Proposed Effect on Developing L2 Pragmatics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticing/Triggering</td>
<td>As learners attempt to produce pragmalinguistic forms (speech acts), they realize that they cannot accurately convey their intended meaning, which triggers their seeking input from others or searching their own developing systems for more appropriate forms.</td>
</tr>
<tr>
<td>Hypothesis Testing</td>
<td>In response to input or feedback that targets their production of speech acts, learners conduct “trial runs” (Swain, 2005) in which they modify their pragmatic output.</td>
</tr>
<tr>
<td>Metalinguistic/Reflective</td>
<td>When learners are required to struggle over the production of pragmalinguistic forms, they use language to reflect on the form and function of the speech acts being attempted.</td>
</tr>
</tbody>
</table>

38
Beginning with Swain’s (1985, 1995) call for SLA researchers to consider the crucial role that Output plays in the acquisition process, studies of the effects of Output have grown in number and influence (e.g., Izumi, 2002; Swain & Lapkin, 2002). Within the realm of L2 pragmatics research, Belz and Kinginger (2003), building on Kinginger’s earlier (2001) research, explored the development in use of terms of address by 14 university-level learners of German who were involved in email interaction with native speakers of German. Adopting a sociocultural theoretic approach, Belz and Kinginger found that the majority (10 of 14) learners exhibited development in the use of address forms over the term of the course, and posited that the opportunities for meaningful interaction with the native speakers contributed to the target language development of the learners.

Similarly, Liddicoat and Crozet (2001), in a case study of ten Australian university-level learners of French, sought to describe how learners could develop the ability to respond appropriately to a question that clearly has different sociopragmatic expectations attaching to it in the two languages/cultures (Australian English and French). The question, “T’as passé un bon week-end?” was presented and the learners were given the opportunity to interact in dyads (in a role-play activity). Before the interaction, however, the learners participated in a four-phase instructional module. By the end of the study, it was apparent that the learners were able to incorporate items related to the content of the appropriate responses, but were not as successful in adapting forms from the acceptable responses. These results appear to support the notion that the output associated with dyadic interaction or small group work (Gass & Varonis, 1985; Long, 1990) facilitates interlanguage pragmatic development, but that there must be an aspect of instruction that targets sociopragmatic development in order to help learners acquire the ability to apply learned pragmatic forms in appropriate contexts.

Martínez-Flor and Fukuya (2005) found that, following six two-hour sessions spread over a sixteen-week semester, both implicit and explicit instructional groups improved in both written and oral production tasks that targeted the use of suggestions. They suggested that the improvement by the group receiving explicit instruction followed from the fact that metapragmatic awareness was raised during the explicit instruction concerning the forms and functions of the suggestions played for them on a videotape. The implicit instruction group, on the other hand, also improved significantly from pretest to posttest. The authors proposed that the improvement of the implicit group was likely a result of their having practiced the forms in role-
plays that were part of the treatment in that study (p. 474). Their findings suggest that Output may be an important factor in the acquisition of pragmatic competence in a second language. It may also be that, as with the Tateyama (2001) study cited earlier, the treatment was spread over time to the extent that differences in instructional effects were not maintained, or possibly not measured immediately for all instructional types.

The findings that Output generally supports IL pragmatic development were reflected in Davies (2004), who took an interactional sociolinguistic approach to the analysis of classroom interaction in a German as a foreign language (GFL) class as well as of materials to be used in GFL instruction to English speakers and ESL/EFL instruction to German speakers. Davies, in a qualitative exploration, noted that where learners and teachers are encouraged to focus on interactional competence, pragmatic development is evident as well. As with the Belz and Kinginger (2003) study, these results again point to a relationship between the active production of pragmalinguistic forms in meaningful interaction and L2 pragmatic acquisition or development. Like Belz and Kinginger, Davies found that the social aspects of language learning, as reflected in her focus on the developing interactional competence of the second language learners involved in the study, can play an important part in promoting the effectiveness of Output as a catalyst for second language acquisition.

A final study reviewed here that has relevance for Output as a factor in developing L2 pragmatics is Bardovi-Harlig and Salsbury (2004). In that longitudinal study of oppositional talk (disagreement), authentic conversational data from 12 participants, all adult ESL learners, interacting with NS graduate students was analyzed for evidence of disagreement and, subsequently, for evidence of development in the use of oppositional talk in pragmatically appropriate ways among the learners over the course of two semesters. It was found that learners with different L2 learning styles (e.g., conservative, risk-taking) showed different developmental patterns. However, there were some general patterns of development to be observed, a fact that has implications for developmental stages of pragmatic acquisition as well as for the effect of Output on developing L2 pragmatics. Among the changes that Bardovi-Harlig and Salsbury observed were that the learners incorporated more postponement of disagreement, elements of humor, and statements of agreement within the oppositional talk as they developed. The researchers proposed that both interaction with NS interlocutors and opportunities to produce
necessary pragmatic forms (Output) were likely contributors to the development of their learners observed in the conversations over time.

The studies reviewed in this section (see Table 5) have each served to clarify a part of the picture of how Output impacts the acquisition of L2 pragmatic competence. The contributions of the various studies, whether controlled or descriptive in design, are useful in the effort to understand developing L2 pragmatics and Output more fully. However, although the studies reviewed here offer some important evidence as to the potential effects of Output on the development of L2 pragmatic competence, they have done little to target more precisely the nature of the contribution of Output to the process of acquiring L2 pragmatic competence. Some questions that remain outstanding are whether passive recognition of pragmatic appropriateness and active production of pragmatic speech acts respond differently to output-based instructional treatments and whether different types of pragmatic speech acts are affected differently by Output. These questions are addressed in the present research effort.

Table 5
Studies on Output and Developing L2 Pragmatics

<table>
<thead>
<tr>
<th>Study</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liddicoat &amp; Crozet (2001)</td>
<td>Learners who produced targeted pragmatic forms in interaction with peers showed development in knowledge, but lacked the sociopragmatic ability to adapt forms to different social contexts.</td>
</tr>
<tr>
<td>Belz &amp; Kinginger (2003)</td>
<td>Most of the university-age GFL learners studied (10 of 14) progressed in the use of address forms following interaction with NSs of German.</td>
</tr>
<tr>
<td>Bardovi-Harlig &amp; Salsbury (2004)</td>
<td>Adult ESL learners (12) were able to develop in the use of pragmatic forms and strategies over time, likely because of opportunities for output and interaction with NSs.</td>
</tr>
<tr>
<td>Davies (2004)</td>
<td>In ethnographic description, GFL learners were observed to progress in the use of pragmatic forms when they were encouraged to focus on aspects of interactional competence.</td>
</tr>
<tr>
<td>Martínez-Flor &amp; Fukuya (2005)</td>
<td>Both implicit and explicit instructional groups improved in written and oral production tasks targeting the use of suggestions; both had opportunities for output in practice.</td>
</tr>
</tbody>
</table>
Limitations of Output-Based Studies for Developing L2 Pragmatics

The studies examined in the previous section represent a subset of L2 pragmatics research that has implications for the effects of Output on developing L2 pragmatic competence. Because the body of research in the area of developing L2 pragmatics has yet to focus specifically on Output as a causative variable, there are several limitations associated with these studies from the perspective of a complete model of L2 pragmatic acquisition. One limitation of the studies reviewed is the fact that many of them have been descriptive and exploratory in design, meaning that the results of the studies that have implications for Output in pragmatic development cannot readily be compared in a meaningful way. Also, because the studies involved interaction of some kind (e.g., Swain & Lapkin, 1995), it is difficult to determine whether the proposed effects of Output are due to output alone or to output in the context of interaction.

Another important limitation of the output-focused studies is that, aside from Martínez-Flor and Fukuya (2005) and, to some extent, Yuan (2001) there has been little effort to contrast written and oral responses to DCT-type tasks within the context of empirical studies. Certainly, several researchers have provided very useful discussions of the merits of either oral or written DCTs in more theoretical terms or in the context of a review of research (e.g., Boxer, 2002; Kasper & Roever, 2005). In the present study, the inclusion of both oral and written DCT instruments should provide an opportunity to gain insight into the important question of whether there is a significant difference in the way these two elicitation instruments indicate developing pragmatic competence.

L2 Pragmatic Speech Act Development

Ohta (2001), in her longitudinal study of learners of JFL, proposed a six-stage developmental sequence for alignment and acknowledgment expressions in Japanese (e.g., Aa soo, desu ka? [‘Oh, really?’]; hai [‘yes’]). Table 6 summarizes Ohta’s findings. The concept of developmental stages in the acquisition of L2 pragmatic competence is significant, not only because it suggests a parallel with general SLA processes (e.g., Pienemann’s, 1989, proposed five stages of developmental readiness in his teachability hypothesis), but also because it may
indicate a differential effect of instruction on different types of pragmatic speech acts. It may be that second language learners must be developmentally ready to acquire the ability to recognize or produce certain types of speech act appropriately. Also, the syntactic differences among different speech act types may be an important factor in establishing an acquisition order.

If the acquisition of second language pragmatic competence is analogous in some ways to the acquisition of L2 morphosyntax, a learner may respond to instruction with respect to speech acts that are typically produced in response to an interlocutor’s proposal or prompt differently than to speech acts that a speaker typically must produce without an interlocutor’s initial cue. In such cases, learners may develop the ability to recognize more “difficult” speech acts before they can accurately produce them. In such circumstances, learners may experience what Ohta referred to as delayed learning. The potential for speech acts of different types to respond differently to a given instructional treatment is one of the research questions addressed in the present study (Research Question 2).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Alignment Expressions</th>
<th>Acknowledgment Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>preformulated questions only</td>
<td>none, except when scripted</td>
</tr>
<tr>
<td>2</td>
<td>used only when scripted</td>
<td>rarely used</td>
</tr>
<tr>
<td>3</td>
<td>used when prompted (some spontaneous)</td>
<td>occasional use increases</td>
</tr>
<tr>
<td>4</td>
<td>used with more facility; expressions beyond Aa soo, desu ka? develop</td>
<td>minimal expressions beyond hai</td>
</tr>
<tr>
<td>5</td>
<td>limited range of expressions modeled by the teacher develops</td>
<td>range of expressions expands</td>
</tr>
<tr>
<td>6</td>
<td>used more spontaneously and appropriately</td>
<td>range of expressions used appropriately</td>
</tr>
</tbody>
</table>

The notion of developmental stages in L2 pragmatics suggested by Ohta (2001) is further supported in Dewaele (2004), who proposed that the ability of a speaker/listener to select from
among possible pronoun choices developed in a three-stage, nonlinear process. Dewaele did not suggest that the stages are invariant or inevitable. In fact, there is support, Dewaele argued, for the notion that “individual paths” (p. 398) to the acquisition of sociolinguistic competence exist. Also, the gradual stabilizing of the pronominal system that is implied in the stages only occurs, it is argued, in the presence of sufficient, frequent use of the L2 to the point that the learner is able to develop the implicit knowledge necessary for that stability. The stages that Dewaele discussed are summarized in Figure 2.

**Figure 2.** Dewaele’s (2004) stages of development toward implicit knowledge.

Another finding that has implications for the notion of a developmental aspect in connection with pragmatic features is that of Beebe and Waring (2004). In a descriptive study of 40 ESL learners at two linguistic proficiency levels, they investigated the ability of learners to use adverbials (e.g., “really,” “never,” “extremely”) to affect the pragmatic tone of an utterance effectively. Employing a six-item DCT, the researchers asked the learners to respond to each situation based on what they would say and, interestingly, what they “would feel like saying” (p. 235). Beebe and Waring found that the lower-level learners were unable to make use of adverbials to convey a range of pragmatic tone as the higher-level learners were.

Finally, another developmental sequence, this one relating to the development of pragmatic strategies, was offered by Bardovi-Harlig and Salsbury (2004) in response to their
analysis of longitudinal data collected from adult ESL learners over about one year. The focus of the study, reviewed earlier, was the production of disagreements, or oppositional talk. The proposed developmental sequence offered by the researchers is summarized in Table 7, adapted from Bardovi-Harlig and Salsbury (2004, p. 218).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong Disagreement</td>
<td>characterized by the frequent use of “no”</td>
</tr>
<tr>
<td>Inclusion of Agreement</td>
<td>attempts to find some agreement within the overall oppositional talk</td>
</tr>
<tr>
<td>Postponement of Disagreement (A)</td>
<td>disagreement elements are delayed, but still introduced within the learner’s turn</td>
</tr>
<tr>
<td>Postponement of Disagreement (B)</td>
<td>introduction of disagreement elements delayed until a later turn within the sequence of turns</td>
</tr>
</tbody>
</table>

**Proficiency and Developing L2 Pragmatic Competence**

Investigations into instructional effects on L2 pragmatic development often touch on the issue of whether pragmalinguistic proficiency and sociopragmatic proficiency are correlated with learners’ general proficiency levels. This issue has relevance for the second research question in the present study concerning the possibility of instruction having differential effects on different types of pragmatic speech acts because of developmental sequences, as suggested in the research of Dewaele (2004) and Ohta (2001). Rose (2000) addressed concerns about different proficiency levels and ages of the learners in his study of Hong Kong elementary students of EFL from three grade levels, with average ages of 7 (n=40), 9 (n=34) and 11 (n=39). The students were divided into two test groups, the first of which (n=15) took a questionnaire but did not participate in data collection, and the second of which (n=98) was subdivided into an English group (n=53) and a Cantonese group (n=15). Pragmalinguistic proficiency, the ability both to choose the correct speech act in a given context and use the correct linguistic means to express it, and
sociopragmatic proficiency, the ability to follow social norms of appropriateness in performing illocutionary acts, were the constructs of interest in Rose’s study.

Rose (2000) asked about the range of pragmalinguistic proficiency in English requests, apologies, and compliment-requests, evidence of pragmalinguistic development among the three groups, as well as the range and development of sociopragmatic proficiency and development among the three groups, finally looking for evidence of influence of Cantonese in the children’s requests, apologies, and compliment responses. Using a cartoon oral production task (COPT) to elicit the targeted speech acts, Rose found that of 524 speech acts, the majority (370 across all groups) were conventional indirect, based on Blum-Kulka, House, & Kasper’s (1989) analytical framework (pp. 39-41). He further found that pragmalinguistic considerations appeared to take precedence over sociopragmatic concerns among these early-stage EFL learners (p. 55). By examining the responses of each of the three distinct age groups, Rose did find some evidence for development of pragmatic competence among the older students, but stipulated that further research among a broader age range (perhaps from ages 7-17) would be needed to understand the possible correlation between age and L2 pragmatic development among this population, and certainly would be needed in order to understand the implications of age for developmental L2 pragmatics among the broader language learning population.

In another study with relevance to the issue of development in L2 pragmatic proficiency, Ravid, Olshtain, and Ze’elon (2003) explored the question of how native speakers of a language modify their speech in terms of both pragmatic and linguistic features when conversing with non-native speakers. Their participants were grade-schoolers in Israel, one group of whom were native speakers of Hebrew, and the other of which were Russian-speaking learners of Hebrew. The researchers found that the native speakers had differing levels of ability to lead a conversation using foreigner talk, and to deal with breakdowns in communication that came up. A logical conclusion from the findings is that individual speaker/listeners, for various possible reasons, have different levels of pragmatic awareness. Such a difference is analogous to the differences in linguistic proficiency often mentioned in the general SLA literature (e.g., Dörnyei & Skehan, 2003).

The relationship between general linguistic proficiency and pragmatic development is an important question with implications for both general SLA and developing L2 pragmatics research. The study of 40 ELLs carried out by Beebe and Waring (2004), reviewed earlier,
examined just this relationship. The researchers’ finding that learners at different levels of L2 proficiency responded differently to a DCT, not only in terms of vocabulary, but also in terms of pragmatic strategies indicating pragmatic tone, led Beebe and Waring to propose that there might be an order of difficulty connected with the acquisition of pragmatic tone. Their speculation was that intense or sarcastic tones might be easier to acquire, whereas such subtle tones as assertiveness might be more difficult to acquire with accuracy.

Because the relationship between general language proficiency and developing pragmatic proficiency in a second language is not entirely clear, it is probably best to control for the possible mitigating effects of one on the other when conducting research. In the case of the present study, the possible effects of general proficiency are accounted for by testing only intermediate-level learners and above. In this way, beginning-level students, who may lack the linguistic resources to adequately convey the sociopragmatic knowledge that they have, are not compared with more advanced learners inappropriately.

Assessing Developing L2 Pragmatic Performance

DCT Rating Scales

Rating scales used to assess L2 pragmatic performance have become more rigorous in recent years (Cohen, 2004). However, such rating scales tend to be impressionistic. Hudson, Detmer, and Brown (1995) utilized a rating scale for analyzing learners’ refusals by rating them in six areas: (a) use of the correct speech act, (b) typicality of expression, (c) appropriateness of amount of information, (d) formality, (e) directness, and (f) politeness. The use of discrete categories for rating the responses is a useful approach that likely enhances the reliability of the resulting scores, but the fact remains that for at least five of the six categories (and possibly all six) the rating or score that results is entirely dependent on the impressions of the rater or raters. The impressionistic nature of L2 pragmatic rating scales is not surprising, given the fact that determinations of pragmatic appropriateness are dependent on context and culturally-dependent background knowledge, and are thus subject to human interpretation.

In addition to the discrete categories approach employed by Hudson, et al. (1995), another approach to rating DCTs is to adopt a holistic rating scale (rubric), with indicators and examples serving to distinguish among the levels of the scale (McNamara, 1996; Weir, 2005).
This approach to assessment is adopted in the present study. Holistic scales, such as those utilized by the Educational Testing Service (ETS) for the new Internet-based TOEFL examination (speaking portion) and the American Council on the Teaching of Foreign Languages (ACTFL; 1986) proficiency guidelines enable raters to consider a response based on its overall acceptability, while also making use of categorical descriptors and, in some cases, examples, in order to clarify the criteria by which levels are differentiated.

**Raters and Rater Training**

In order to obtain valid scores for DCT responses, it is essential that researchers recruit raters with characteristics that fit the research questions addressed and that those raters be trained sufficiently. Brown (2004), in a discourse analysis-based study of oral interview tests, found that the choice of interviewer has important implications for learners’ test performance. The situation with regard to recruiting raters for scoring pragmatic responses is analogous. McNamara (1996) reported that raters may exhibit variability in one or more of several areas, including overall leniency toward responses, harshness or leniency toward particular groups of test-takers, interpretations of the rating scale used, and consistency of ratings, among others (pp. 123-124).

Weir (2005) defined rater training as “a systematic process to train raters to apply the rating scale and the mark scheme in a consistent way” (p. 190). Rater training ostensibly mitigates the influence of some of the differences that exist among raters. If non-experts are employed as raters (as is the case in the present study), training and familiarization with the rating scale is particularly important. McNamara (1996) reported that rater training has been found to be as significant a factor as rater experience in terms of consistent, reliable marking or scoring. The program of rater recruiting and training utilized in the present research effort is detailed in the next chapter. At this point, what may be stated with some confidence is that a careful assessment of the validity and reliability of the DCT scoring rubric proposed for the study is warranted by a survey of the language testing literature. Such an assessment is the focus of the third research question.

**Summary**

Based on the limitations revealed in a review of the research concerning the effects of Output on developing L2 pragmatics, there is a theoretically motivated need to investigate the
effects of Output in instruction on the acquisition of second language pragmatic competence. Also, evidence for possible developmental stages in the acquisition of L2 pragmatics suggests that there may be a differential effect of instruction on the acquisition of speech acts of different types. Additionally, a brief examination of some research that has touched on the issue of general language proficiency as it relates to L2 pragmatic development suggests that it may be best to control for, or at least account for, the possible effects of one on the other when conducting research. Finally, a consideration of research that has focused on the validation of language testing as it relates to the assessment of L2 pragmatic development indicates that there is a need to carefully assess the reliability and validity of the DCT rating scale (rubric) utilized in the present study.
Kasper and Rose (2002) have proposed that acquisitional pragmatics research needs to take task type into consideration, make use of sufficiently detailed transcriptions of the type used in conversation analytic research (e.g., Kasper, 2004a; Markee, 2000, 2004), look at individual learner contributions in the context of interaction, and present adequately detailed and contextualized data excerpts when discussing findings. With Kasper and Rose’s challenge in mind, the third chapter details the research design, participants, materials, procedure, and scoring and analyses associated with the present study. A review of the pilot study, conducted several months prior to the principal study, follows.

**The Pilot Study**

A small-scale pilot study of Output in the context of video-based pragmatic instruction and its possible effects on developing L2 pragmatic competence was carried out some five to six months prior to the principal research study. The following sections summarize the major elements of study design and results from the pilot study, along with a discussion of the limitations of the pilot project and several suggestions for the principal study in response to those limitations.

**Pilot Study Participants**

Nine adult English as a second language (ESL) students from an intact ESL class, ranging in age from 18 to 40, participated in the study. The participants represented six different first language (L1) backgrounds: French, Japanese, Korean, Mongolian, Spanish, and Turkish. There were five male and four female participants, all of whom were of intermediate to high-intermediate proficiency, based on their placement at the intensive English program (IEP) at which they studied. None of the participants in the pilot study participates in the present study. The researcher was also the teacher of the participants.
Pilot Study Materials

The central instructional tool was a pragmatics-focused instructional video developed and produced by the researcher. The video, based loosely on Mir’s (2001) proposed model of video-based pragmatic instruction for learners of Spanish, consists of seven short vignettes that portray realistic interactions between native speakers of English making use of one or more pragmatic strategies and the associated pragmalinguistic forms. For the pilot study, the vignettes of interest were those that portrayed request refusals (4 vignettes) and compliment responses (3 vignettes).

Pilot Study Procedure

The central activity of the study was video-based instruction designed to raise the awareness of pragmatic strategies and associated pragmalinguistic forms. Bou-Franch and Garcés-Conejos (2003) advocated an approach to teaching pragmatics in the classroom that enhances learners’ sociopragmatic and pragmalinguistic knowledge by not only exposing them to input in the L2, but also by explicitly teaching linguistic politeness. The pilot study treatment and tasks loosely followed their model of instruction, which was informed by Brown and Levinson (1987) and Scollon and Scollon (1995). However, in the pilot study, formal instruction was replaced by an output-focused activity (negotiated interaction).

The nine participants were divided into two groups of six and three. Originally, groups of six and five were formed, but two students failed to complete all of the tasks, and both of those participants represented the second, already smaller group. Both groups completed a six-item written discourse completion task (DCT; Beebe, Takahashi, & Uliss-Weltz, 1990; Lorenzo-Dus, 2001; Liao & Bresnahan, 1996), which functioned as a pretest (see Figure 3).

You are a junior in college. You attend classes regularly and take good notes. Your classmate often misses class and asks you for the lecture notes.

Classmate: Oh no! We have an exam tomorrow but I don’t have the notes from last week. I am sorry to ask you this, but could you please lend me your notes again?

You: ______________________________________________________________
     ______________________________________________________________

Classmate: Oh, well I guess I’ll have to ask somebody else.

Figure 3. Sample pilot test DCT item.
Following the initial (pretest) DCT, the participants all viewed the seven video vignettes on two successive class days as part of their classroom instruction. The participants were encouraged to take notes during the viewing sessions. Following Izumi (2002), the Output group participants were advised to take notes in order to be able to reconstruct the dialogues, whereas the non-Output participants were instructed to take notes in order to be able to answer comprehension questions about them.

Next, the Output group (participant $n = 6$) was tasked with an output-focused task: each participant was given a copy of the transcript of the first video vignette and was given about one minute to take notes based on the transcript. Then, the learners were paired and, using their notes, worked to reconstruct the original dialogue as closely as possible in dyads. The process was repeated for each of the other six vignettes over the course of two days.

The three participants in the non-Output group, serving as a control group, were given a comprehension task, and worked individually to complete their task. Following the video viewing, each student was provided a copy of the transcript of the first vignette. Just as with the Output group, each participant had approximately one minute to take notes from the transcript. After the note taking, each individual was given a set of comprehension questions. An example question is provided in Figure 4, corresponding to the transcript portion above it:

R: Hey, what did he say about the homework assignment?

M: He said we have an extra day to finish and turn it in.

(1) How much extra time do the students have to complete their homework assignment?

   a. Two days   b. one week   c. one day   d. Two hours

*Figure 4. Sample pilot test comprehension question*
Pilot Study Scoring and Analysis

The instrument. The test of development in pragmatic competence is performance on the pretest and posttest written DCT. A significant difference in pretest to posttest scores may be interpreted as learning in the area of interlanguage (IL) pragmatics. Essentially, the DCT functions to create a scenario to which the participant/informant must respond. Typically, there is an initial statement outlining the context within which a dialogue occurs. Next, the first line of the dialogue is presented, and the participant is then given an opportunity to respond in the way that she believes most appropriate.

Scoring. The participants’ responses to the DCT items were rated on a four-point scale:

(1) response is unacceptable pragmatically given the context
(2) response is generally unacceptable pragmatically in this context, though perhaps not in all contexts
(3) response is generally appropriate given the context, but contains one or more noticeable pragmatic flaws that affect the intended meaning
(4) response is acceptable pragmatically given the context, not noticeably affected by any errors

There were two raters participating in the scoring process: the researcher and a colleague, also a native speaker of English and an experienced ESL instructor. The four-point rating scale was discussed thoroughly by the two raters and then both raters scored one-third of the responses so that ratings could be compared in order to promote a high level of consistency and reliability. The remaining two-thirds of the responses were scored individually. After comparing the item scores that were rated together, it was determined that the raters agreed on 94.1% of the items rated together (109 compared with 119 total points given for 42 items, with a maximum of 168 total points possible), and the internal consistency and reliability of the rubric was calculated using reliability analysis in the SPSS statistical package at .81 (standardized item alpha).

Analysis of Pilot Study Data

Statistical analyses carried out to compare the performances of the two groups were a repeated measures ANOVA to compare pretest and posttest scores, as well as multiple t-tests for multiple comparisons in order to explore any interaction that is revealed by the ANOVA. These tests were determined to be most appropriate because the study sought to compare two groups, each with the between-subjects factor of group (+ Output or – Output) and the within-subjects
factor of accuracy of pragmalinguistic production (pretest to posttest performance difference on the DCT).

**Pilot Study Results**

The total number of item responses analyzed was 126 (item \( n = 63 \) each for the pretest and posttest DCT). Table 14 depicts the mean pragmalinguistic production DCT scores.

### Table 8

*Mean Pilot DCT Scores by Output Group Condition*

<table>
<thead>
<tr>
<th>Group</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Gain Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( N )</td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td>+ Output</td>
<td>42</td>
<td>2.71</td>
<td>.790</td>
</tr>
<tr>
<td>- Output</td>
<td>21</td>
<td>2.81</td>
<td>.512</td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>2.75</td>
<td>.706</td>
</tr>
</tbody>
</table>

Note: maximum score for each item was 4.

The repeated measures ANOVA revealed a significant increase in scores from pretest to posttest, \( F(1, 61) = 7.73, p < .05 \), and a significant Group x Time interaction, \( F(1, 61) = 7.73, p < .05 \). In order to explore the group effects, paired samples \( t \)-tests were conducted. Paired samples \( t \)-tests with adjusted alpha-levels revealed: (a) a significant increase from pretest to posttest DCT scores for the + Output group, \( t(41) = -4.60, p < .001, d = 0.89 \), two-tailed; and (b) no significant increase from pretest to posttest for the – Output group. In addition, independent samples \( t \)-tests confirmed no significant difference between pretest scores for the two groups, but a significant difference for the posttest scores, \( t(61) = -3.35, p = .001, d = 0.93 \), two-tailed.

The ANOVA procedure, of course, assumes a normal distribution of the data, and with the small sample size of the proposed study in terms of number of participants, employing the ANOVA represents a limitation of the pilot study. Also, the results of the \( t \)-test for comparing the group means must be interpreted cautiously for similar reasons. Nevertheless, the pilot study results suggest some type of effect for Output on the acquisition of certain L2 pragmalinguistic forms.
Discussion and Limitations of the Pilot Study

The results of the pilot study generally are encouraging, in that they suggest a possible beneficial effect for Output in the context of instruction on pragmalinguistic forms. There are a number of limitations of the pilot, however, that need to be acknowledged. Recalling that the purpose of carrying out the pilot study was to become familiar with the data collection methods and to clarify some of the constructs likely to be part of the principal research study, it is important to note that the limitations discussed here are best viewed in terms of their potential to benefit the principal research study. For each limitation or drawback presented below, suggestions for how to apply the lessons learned to the principal study are offered.

The first limitation is the small participant sample size, which led to a shift from the learner as the unit of analysis to the DCT response as the unit of analysis. In order to obtain meaningful results from the pilot, the statistical analyses (the ANOVA and \( t \)-tests) were conducted based on the individual item responses on each participant’s DCT. The low participant \( n \) for both groups of learners in the pilot study means that there is a high potential for sampling error in the pilot study results, leading to conclusions that may not reflect the population at large (Bachman, 2004; Creswell, 2005). To promote increased validity in the principal study, the largest sample possible will be drawn from the available ESL learner population. In this way, the learner can be maintained as the unit of analysis and the results can reflect the population of adult ESL learners to a much more satisfactory extent.

A second important limitation of the pilot study is that the output-oriented treatment involved dyadic interaction, which means that rather than simply measuring the effects of Output on the learners’ DCT responses, the variable of interaction was also being measured inadvertently. In response, the Output treatment task in the principal study follows Izumi (2002) more closely in utilizing text reconstruction tasks that do not involve interaction with a classmate. The final pilot study limitation discussed here is the lack of an oral output measure in the tasks. As Yuan (2001) has noted, oral production tasks can elicit much more useful data than corresponding written tasks, so the principal research study benefited from the incorporation of a relatively more interactive and spontaneous oral DCT instrument that provided the participants with a situation and prompt and then gave them the opportunity to respond by speaking a pragmatically appropriate response.
The Principal Research Study

The principal study built upon the pilot study described above, benefiting from both the strengths and limitations of that smaller-scale effort. The principal study featured a pretest-posttest design in order to address three primary research questions, two of which are further divided into three sub-questions. To begin this discussion of the present study, a schematic representation of the research questions, along with the statistical analyses used to address each question, follows in Table 9. As indicated in the table, the first two research questions were addressed by testing the + Output and – Output group participants’ pretest-posttest performances utilizing repeated measures ANOVA and follow-up t-tests as well as the calculation of effect sizes. The third research question, related to the reliability of the rating rubric as utilized by the non-expert raters, was addressed by calculating Cronbach’s alpha (or coefficient alpha).

Table 9
Summary of Research Questions

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Analyses</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Effect for Output on developing L2 pragmatic competence?</td>
<td>repeated measures ANOVA analyses conducted for each of the three outcome measures; follow-up t-tests conducted in order to determine the source of any interaction identified by the ANOVA; calculation of effect size</td>
<td>Bachman (2004); Cohen (1988); Izumi (2002); Mackey &amp; Gass (2005)</td>
</tr>
<tr>
<td>(1a) PAJT outcome?</td>
<td></td>
<td>Garcia (2004)</td>
</tr>
<tr>
<td>(1b) WDCT outcome?</td>
<td></td>
<td>Beebe, et al. (1990)</td>
</tr>
<tr>
<td>(1c) ODCT outcome?</td>
<td></td>
<td>Yuan (2001)</td>
</tr>
<tr>
<td>(2) Differential effect of Output on two speech act types?</td>
<td>repeated measures ANOVA and follow-up t-tests carried out for the two speech act types; calculation of effect sizes</td>
<td>Bardovi-Harlig &amp; Salsbury (2004); Cohen (1988); Dewaele (2004); Ohta (2001)</td>
</tr>
<tr>
<td>(3) Rubric reliable and valid?</td>
<td>Cronbach’s alpha calculated to determine the rubric’s internal consistency and reliability</td>
<td>Bachman (2004); Cronbach (1984); Weir (2005); Yamashita (1996)</td>
</tr>
</tbody>
</table>
Research Design

As mentioned, the principal study built and expanded upon the smaller pilot study. The overall design of the study benefited from some limitations observed in the pilot study, and drew on traditions of research carried out within the framework of Swain’s (1985, 1995, 1998, 2005) Output Hypothesis. The influence of previous studies in the area of developing L2 pragmatics is also evident in the design of the present study. The focus of the study remained the effect of Output on adult ESL learners’ developing pragmalinguistic proficiency with respect to two different types of pragmatic speech acts (the first two research questions). Consideration was also given to the reliability and validity of the proposed rubric for assessing the developing interlanguage pragmatic production of ESL learners. The research design is depicted in Figure 5.

![Research Design Diagram]

* Treated in a separate analysis of the raters’ scores for the DCT instruments

*Figure 5. Overview of Research Design.*
Description of the Variables

Independent variable. The independent variable in the study is Output group (two levels: those receiving opportunities for output and those without output), which was explored in the first two research questions.

Dependent variables. Although the overall conceptual dependent variable is developing L2 pragmatic competence, the specific dependent variables (outcome measures) in the study are the judgment of acceptable pragmalinguistic forms (two categories: directives, expressives; pretest-posttest); the oral production of pragmalinguistic forms (two categories: directives, expressives; pretest-posttest); and the written production of pragmalinguistic forms (two categories: directives, expressives; pretest-posttest). The theoretical difference between expressives and directives stems from the identification of a possible developmental sequence in the acquisition of L2 pragmatic competence (Dewaele, 2004; Ohta, 2001). Operationalization of the variables is treated further in the description of the study procedure. The variables in the study (not including the test for reliability of the DCT rating rubric) are summarized in Table 10 below.

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group (+ Output, - Output)</td>
<td>Judgment of Acceptable Pragmalinguistic Forms (Directives, Expressives) from pretest-posttest</td>
</tr>
<tr>
<td></td>
<td>Written Production of Pragmalinguistic Forms (Directives, Expressives) from pretest-posttest</td>
</tr>
<tr>
<td></td>
<td>Oral Production of Pragmalinguistic Forms (Directives, Expressives) from pretest-posttest</td>
</tr>
</tbody>
</table>

Speech acts. The illocutionary force of an utterance, i.e., the intended effect of one’s output on an interlocutor, is tied to the type of speech act that is being performed. There is
precedence in the literature (e.g., Wee, 2004) to consider speech acts to be a subset of 
communicative acts, with the former indicating verbal or linguistic communication and the latter 
including non-verbal signals as well. In that sense, communicative acts might include gestures 
and other paralinguistic acts that are commonly regarded as contributing to a more general 
“intercultural competence” (Molinsky, Krabbenhof, Ambady, & Choi, 2005). The focus in the 
present study was speech act perception and production as indications of developing L2 
pragmatic competence.

Four speech acts were tested in the study (see Table 11): (a) request refusals (Bardovi-
Harlig & Hartford, 1991; Liao & Bresnahan, 1996), (b) compliment responses (Lorenzo-Dus, 
2001; Golato, 2003), (c) advice giving (Matsumura, 2003), and (d) invitations (Beebe, 
Takahashi, & Uliss-Weltz, 1990; Ochs, 1996). The first two speech acts were also tested in the 
ellier pilot study. The two additional speech acts tested in the principal research study fall 
within the general pragmatic category of directives (see Table 2 in the previous chapter). The 
newly added illocutionary pragmatic acts were chosen because, like the two speech acts utilized 
in the pilot study, they have consistent representation in the literature (e.g., Blum-Kulka, House, 
& Kasper, 1989; the DCT used and published by Beebe, Takahashi, & Uliss-Weltz, 1990). These 
four speech acts were also found to be high in frequency in casually observed conversations and 
in the Michigan corpus of academic spoken English (MICASE; Simpson, Briggs, Ovens, & 
Swales, 2002). In addition, compliment responses are interesting from the perspective of 
conversational maxims; Yoon (1991) pointed out that in order to respond to a compliment in a 
pragmatically acceptable manner, one must walk a fine line between Leech’s (1983) politeness 
maxims of modesty (not inflating one’s own status) and agreement (the tendency to agree with 
what one’s interlocutor says). Finally, the ability to group the speech acts tested into categorical 
pairs facilitated the exploration of the second research question of the study.

The choice of these two specific speech act categories (expressives and directives) was 
motivated by the adequate representation of both in the existing pragmatic literature and the 
difference, proposed by the present researcher, between the kind of pragmatic awareness needed 
to carry out each type of speech act. The fact that expressives require the surface realization of 
relatively deep speaker motives whereas directives require the socially acceptable use of 
politeness techniques may mean that a learner at a given stage of L2 pragmatic development can 
successfully recognize and produce one of the two pragmatic speech act types even when they

59
are unable to recognize or produce the other successfully.

Table 11

<table>
<thead>
<tr>
<th>Summary of Speech Acts Tested</th>
<th>Expressives</th>
<th>Directives</th>
</tr>
</thead>
<tbody>
<tr>
<td>request refusals (e.g., “Sorry, I can’t…”)</td>
<td>advice giving (e.g., “You really should…”)</td>
<td></td>
</tr>
<tr>
<td>compliment responses (e.g., “Thanks, I just…”)</td>
<td>invitation (e.g., “Would you like to…?”)</td>
<td></td>
</tr>
</tbody>
</table>

Participants

Adult ESL Learners

A useful distinction when discussing adult learners of English as a second language is between those adult learners who are studying English in the context of high-school or community college-level programs and those who are studying ESL at the university level (international university students) or in intensive English programs (IEPs). Murray (2005) discusses adult ESL learners who immigrate to other nations as political or economic refugees as having special needs that are different from those of international university or IEP students. Most notably, adult refugees may have varying levels of established literacy in their first language (L1), whereas international university students or adult students attending IEPs typically have consistently high levels of L1 literacy.

For the purposes of this study, the population of adult ESL learners investigated was the subset of adult ESL learners attending an IEP in an English-speaking country. For these learners, critical needs include the development of academic reading and writing skills in English and the ability to communicate with English-speaking professors and classmates, which requires the development of pragmatic competence in American English. Research by Davies and Tyler (2005), among others, has indicated the need for even advanced-level ESL learners in post-secondary academic settings such as the U.S. university classroom to develop in the appropriate
use of discourse strategies and understanding of target-culture expectations, particularly during oral interaction with native speakers. Because ESL learners entering post-secondary educational settings in the United States and elsewhere may focus most of their efforts on developing linguistic proficiency for reading and writing and building academic vocabulary (Carkin, 2005), the acquisition of L2 pragmatic competence needs to be understood more fully in order for these learners to be able to adjust to and communicate successfully in the U.S. university setting, whether as students, graduate teaching assistants, or in other roles.

**Study Participants**

The study drew from the population of adult ESL learners at an intensive English program (IEP) at a major southeastern U.S. university. Four different intact ESL classes in two separate rounds of data collection and testing were selected for inclusion in the research, and the final number of participants ($n$) was 34. The use of intact classes for the study was consistent with the ESL classroom focus of the present study. Among the participants, 18 male learners and 16 female learners were selected, and the English proficiency levels of the participants varied from intermediate to advanced, based on their TOEFL scores and placement at the IEP. Generally, the placement levels at the IEP corresponded to approximate TOEFL score ranges such that the intermediate-level learners’ Paper-based TOEFL scores typically were in the range of 450-500, and the advanced-level learners typically had earned Paper-based scores of 500 and upward (or the equivalent on the Computer-based or Internet-based TOEFL). For the present study, the participants’ self-reported TOEFL scores ranged from a 56 Internet-based score (approximately equivalent to a 483 paper-based test score) at the lower end to a 240 Computer-based score (approximately equivalent to a 587 Paper-based score) at the upper end (score conversions based on Educational Testing Service, 2006). Elementary (beginning) level learners were not included in the present study because of concern that their limited English proficiency might prevent them from comprehending all of the materials used in the study.

The first language backgrounds of the study participants as a whole included Arabic, Chinese, French, Italian, Japanese, Korean, Portuguese, Spanish, Vietnamese and Turkish. Table 12 offers an overview of the study participants in the + Output group. As the summary in the table makes clear, the + Output group included 12 male participants and 5 females. Also, there were 10 intermediate-level learners in the + Output group and 7 advanced-level learners. The learners in the + Output group represented 8 different first language backgrounds, including
Table 12
+ Output Group Study Participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>First Language</th>
<th>English Proficiency Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM1</td>
<td>Male</td>
<td>Arabic</td>
<td>Intermediate</td>
</tr>
<tr>
<td>AM2</td>
<td>Male</td>
<td>Arabic</td>
<td>Intermediate</td>
</tr>
<tr>
<td>AM3</td>
<td>Male</td>
<td>Arabic</td>
<td>Advanced</td>
</tr>
<tr>
<td>AM4</td>
<td>Male</td>
<td>Arabic</td>
<td>Advanced</td>
</tr>
<tr>
<td>AM5</td>
<td>Male</td>
<td>Arabic</td>
<td>Advanced</td>
</tr>
<tr>
<td>CF1</td>
<td>Female</td>
<td>Mandarin Chinese</td>
<td>Intermediate</td>
</tr>
<tr>
<td>CF2</td>
<td>Female</td>
<td>Mandarin Chinese</td>
<td>Intermediate</td>
</tr>
<tr>
<td>CF3</td>
<td>Female</td>
<td>Mandarin Chinese</td>
<td>Intermediate</td>
</tr>
<tr>
<td>FM1</td>
<td>Male</td>
<td>French</td>
<td>Advanced</td>
</tr>
<tr>
<td>IM1</td>
<td>Male</td>
<td>Italian</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KM1</td>
<td>Male</td>
<td>Korean</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KM2</td>
<td>Male</td>
<td>Korean</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KM3</td>
<td>Male</td>
<td>Korean</td>
<td>Advanced</td>
</tr>
<tr>
<td>PF1</td>
<td>Female</td>
<td>Brazilian Portuguese</td>
<td>Advanced</td>
</tr>
<tr>
<td>SF1</td>
<td>Female</td>
<td>Spanish</td>
<td>Intermediate</td>
</tr>
<tr>
<td>SM1</td>
<td>Male</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>VM1</td>
<td>Male</td>
<td>Vietnamese</td>
<td>Intermediate</td>
</tr>
</tbody>
</table>

Table 13 provides information on the participants assigned to the – Output treatment group in the study. The summary in the table indicates that the – Output group included 11 female learners and 6 male participants. In addition, as with the + Output group, there were 10
intermediate-level learners and 7 learners at the advanced English proficiency level. The – Output group included learners from 7 different first language backgrounds.

Table 13
- *Output Group Study Participants*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>First Language</th>
<th>Proficiency Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF1</td>
<td>Female</td>
<td>Arabic</td>
<td>Intermediate</td>
</tr>
<tr>
<td>CF4</td>
<td>Female</td>
<td>Mandarin Chinese</td>
<td>Intermediate</td>
</tr>
<tr>
<td>GF1</td>
<td>Female</td>
<td>German</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KF1</td>
<td>Female</td>
<td>Korean</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KF2</td>
<td>Female</td>
<td>Korean</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KF3</td>
<td>Female</td>
<td>Korean</td>
<td>Advanced</td>
</tr>
<tr>
<td>KM4</td>
<td>Male</td>
<td>Korean</td>
<td>Intermediate</td>
</tr>
<tr>
<td>KM5</td>
<td>Male</td>
<td>Korean</td>
<td>Advanced</td>
</tr>
<tr>
<td>PF2</td>
<td>Female</td>
<td>Brazilian Portuguese</td>
<td>Intermediate</td>
</tr>
<tr>
<td>SF2</td>
<td>Female</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>SF3</td>
<td>Female</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>SF4</td>
<td>Female</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>SM2</td>
<td>Male</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>SM3</td>
<td>Male</td>
<td>Spanish</td>
<td>Advanced</td>
</tr>
<tr>
<td>SM4</td>
<td>Male</td>
<td>Spanish</td>
<td>Intermediate</td>
</tr>
<tr>
<td>TF1</td>
<td>Female</td>
<td>Turkish</td>
<td>Intermediate</td>
</tr>
<tr>
<td>TM1</td>
<td>Male</td>
<td>Turkish</td>
<td>Intermediate</td>
</tr>
</tbody>
</table>

The two instructional treatment groups each represented two different intact ESL classes, one advanced-level class and one intermediate-level. As Tables 12 and 13 reveal, the two groups...
of participants represented a sufficiently diverse set of ESL learners to warrant the carrying out of tests of probability such as the ANOVA.

Materials

Video-based Instructional Materials

The video-based material used for the study expanded on the seven-vignette video that was produced by the researcher for the pilot study. Following Mir (2001) and Nedashkivska (2004), the scripts for the videos were designed to reflect actual language use as far as possible. That is, idealized forms of speech acts were not sought, but rather the realizations of speech acts as they naturally occurred as the speakers being filmed attempted illocutionary moves in conversations. The language of academic and social life related to the university experience in the United States was the focus of the videos for instruction. To that end, in some cases the outlines for the vignettes were drawn from the Michigan corpus of academic spoken English (MICASE; Simpson, Briggs, Ovens, & Swales, 2002) and in other cases, from informal observations by the researcher of common conversations that took place in a university setting.

The actors used in the production of the video vignettes were primarily graduate students and ESL instructors. To ensure that the pragmatic norms of American English were reflected in the vignettes, all of the actors were native speakers of American English. Feak and Salehzadeh (2001) noted that video-based input must be carefully utilized to ensure that the visual dimension of the video material complements and contributes to the audio input that it accompanies, rather than distracting from it. With that admonition in mind, the use of video-based materials in the present study was judged to be of great value because of the close relationship between successful pragmatic effectiveness and the appropriate implementation of visual cues such as gestures, facial expressions, and other paralinguistic phenomena.

In all cases, the outlines for the vignettes were relatively unconstrained (i.e., the actors were given the context, the desired speech act, and some key elements that needed to be present) so that as much of the dialogue on the videos as possible reflected “real speech” as used by native speakers of American English. The filming and production of the video vignettes was done by the researcher with some assistance. The pragmatic vignettes, most of which were less than one minute in length, were designed to provide the learners with the opportunity to see and
hear the targeted speech acts in context. Figure 6 provides a sample transcript of one of the pragmatic vignettes.

<table>
<thead>
<tr>
<th>Classroom Request Refusal Vignette</th>
</tr>
</thead>
<tbody>
<tr>
<td>K: Okay everyone, would you please open your books to page 37? Let’s do the questions on page 37, please.</td>
</tr>
<tr>
<td>D: Dr. Kott?</td>
</tr>
<tr>
<td>K: Deborah, yes.</td>
</tr>
<tr>
<td>D: Yes, I was wondering, could we review the old questions before we start the new questions?</td>
</tr>
<tr>
<td>K: Oh, well, that would be a very good idea, but I think I gave everyone two days to finish that assignment, and you may have finished, but the others perhaps are still working through it, so maybe it would be better to wait until tomorrow for that.</td>
</tr>
<tr>
<td>D: I had just, I forgot you gave us an extra day.</td>
</tr>
<tr>
<td>K: Oh, well, please, Deborah, please, could you begin on page 37 for us then?</td>
</tr>
<tr>
<td>D: Sure.</td>
</tr>
</tbody>
</table>

*Figure 6. Sample pragmatic video vignette.*

The task materials for the study were varied in nature, encompassing the four speech acts described earlier, divided into two broad categories. In addition to the expressive speech acts (compliment responses and request refusals) featured in the pilot study, the principal study targeted the directive speech acts of advice giving (Matsumura, 2003) and inviting (Ochs, 1996). In all, there were three vignettes for each speech act, for a total of 12 vignettes. The vignettes used in the study are transcribed in detail in Appendix I.

**Test Instruments**

The testing materials included an instrument to gauge pragmatic perception called a pragmatic acceptability judgment task (PAJT; Garcia, 2004); a written discourse completion task (WDCT; Beebe, Takahashi, & Uliss-Weltz, 1990; Liao & Bresnahan, 1996; Lorenzo-Dus, 2001); and, in addition to that off-line measure, a relatively more realistic interview-style oral production task called an oral discourse completion task (ODCT; Martínez-Flor & Fukuya, 2005). These instruments are discussed in greater detail below.
Pragmatic acceptability judgment task (PAJT). The pragmatic acceptability judgment task (PAJT) is a closed (i.e., multiple choice), twelve-item instrument that gives participants the opportunity to demonstrate the degree to which they can identify the pragmatic intentions and acceptable variants of participants in twelve different dialogues (three for each of the targeted speech acts in the present study). The PAJT was the first measure (pretest and posttest) of L2 pragmatic development completed by the study participants, as it was the least open-ended of the three measures employed in the study (see Appendices A and B). Garcia (2004) focused on the identification of speaker intention in the judgment task in that study. In the present study, the target is both identification of speaker intention and the acceptability of alternative pragmalinguistic forms. For that reason, the PAJT used here asked learners to choose from among possible pragmatic forms that are more or less acceptable in given situations.

Written discourse completion task (WDCT). The next test of development in pragmatic competence was performance on the twelve-item pretest and posttest written DCT (see Appendices C and D). A significant difference in pretest to posttest scores would be interpreted as development in the area of interlanguage (IL) pragmatics. Essentially, a DCT functions to create a scenario to which a participant or informant must respond. Typically, there is an initial statement outlining the context within which a dialogue occurs. Next, the first line of the dialogue is presented, and the participant is then given an opportunity to respond in the way that she believes most appropriate.

Boxer (2002) notes that at this point, DCTs may take one of two forms: open-ended questionnaires simply ask for the subject to supply the relevant speech act in response to the stimulus (the first line); closed questionnaires elicit a speech act from the subject, but then follow the blank line with a reply from the first speaker, so that the response must take into account not only the initial statement, but also the reply to follow (p. 15). Recognizing the limitations of DCT tasks is important (Beebe & Cummings, 1996; Kasper & Roever, 2005). Among the most common criticisms of the written DCT are that the written format may not accurately reflect oral (spoken) language production, and that the task is rather artificial, reflecting what a learner believes she would say, rather than demonstrating what she actually says in conversation. Also, Watts (2003) has reported that written DCTs are particularly susceptible to manipulation by informants, resulting in responses that they would never use in actual discourse. On the other hand, Chaudron (2003) suggested that DCTs may not allow learners to sufficiently express their
pragmatic competence, particularly in the sense that learners are bound to some extent by the situations presented in the DCT. In order to address such concerns, the design of the written and oral DCTs utilized for the present study included a variety of academically oriented contexts and every effort was made to encourage the participants to provide realistic responses.

Further concerns about the validity and use of DCTs in non-western cultural contexts have been raised by Rose (1994). Additionally, Sasaki (1998) reported on research that DCTs do not correlate with some other measures of pragmatic production, such as role-play tasks. In a study of compliment responses, Golato (2003) examined the differences between DCTs and naturally occurring talk. Despite these limitations and concerns, the written DCT does provide a measure of pragmatic competence and intuition that can be quantified and compared more easily than many other measures. For the present study, the written and oral DCTs utilized are adaptations of instruments used in previous L2 pragmatics research (e.g., Beebe, Takahashi, & Uliss-Weltz, 1990; Sasaki, 1998; Yamashita, 1996). Adapting previously used instruments promotes validity and reliability of the DCT instruments and helps to make the results of the present study more comparable to previous findings.

Oral discourse completion task (ODCT). Yuan (2001), in an investigation into the relative benefits and limitations of various methods of empirical data collection used in pragmatics research, reported that oral DCTs generally mirror both the benefits and limitations of the written DCT, with the important exception that the oral DCT tends to elicit “a significantly larger number of natural speech features than the written DCT does” (p. 271). For the present study, the twelve-item oral DCT was conducted in an interview-style format, in that the participants were exposed to a situation and a prompt, to which they were instructed to provide an acceptable spoken response (Martinez-Flor & Fukuya, 2005).

For the sake of practicality and ease of data collection and analysis, the interview questions in the ODCT task were written out and the procedure consisted of the participant reading the prompt and supplying an oral response, which was recorded for later analysis (Martinez-Flor & Fukuya, 2005). For the expressive speech acts (request refusals and compliment responses), the participants were expected to react to the request or compliment provided by the prompt. In the case of the directives, the participants were asked to read the description of the situation, then to provide either an appropriate piece of advice or an appropriate invitation in response. In order to control for the potential bias of gender and social
status (Bachman, 2004), the prompts and situations presented to the participants in both the
WDCT and ODCT featured interlocutors of varying relative social status (e.g., classmates,
professors, friends, parents) and both genders (see Appendices E and F).

Procedure

Group Assignment

As indicated in the discussion of participants, the learners were assigned to one of two
treatment groups: those receiving opportunities for output (the + Output group) and those who
did not (the – Output group). The learners were tested in their intact ESL classes, and two classes
were selected for the study in each of two rounds of data collection: one class at the intermediate
level and the other at the advanced level. The assignment of each class to one of the two Output
groups was random in the first round of data collection, with the advanced class taking the +
Output assignment, and the intermediate group assigned to the – Output status for round one. In
the second round of data collection, the class assignments were reversed, so that the intermediate
class was assigned to the + Output treatment, with the advanced class, - Output. The two sessions
of data collection took place approximately 14 weeks apart, and the procedure for data collection
was identical in each session. That is, after the fifth and final day of data collection and posttests
for the first round, about 14 weeks elapsed before the first day of pretests for the second round.

Informed Consent and Pretesting

Day 1 of the study began with the distribution of informed consent forms to both the
+ Output and – Output groups (see Appendix L). Next, each group was instructed as to the basic
procedure that would be followed over the course of the week. Specifically, both groups of
learners were told that they would complete three pretests, then view video vignettes and have
the opportunity to complete a learning activity after each vignette. The learners in the + Output
group were informed that they would be asked to reconstruct the vignettes after viewing them
and having access to the transcripts for a limited period. It was explained to them that they would
sometimes be reconstructing the vignettes in writing, and at other times, orally. The – Output
group participants were told that they would be asked to respond to several comprehension and
extension questions after viewing each video vignette and having access to the transcript briefly.
Finally, all participants were told that they would be taking three posttests that corresponded to
the pretests at the conclusion of the week. The participants were not informed of the pragmatic focus of the study, nor were they told about the treatment activities to be completed by the other group.

Following the completion of the informed consent forms and instructions, the participants were given the three pretest instruments to complete. First, each learner completed the PAJT by choosing their responses on the papers provided to them. Next, the learners completed the WDCT by writing out their responses to the situations and prompts on the instrument. Finally, each learner completed the ODCT by reading the situations and prompts on that instrument and recording their spoken responses on either personal computers (utilizing Microsoft® Sound Recorder, version 5.1) or on handheld digital audio voice recorders (Olympus® WS-200S and DS-2 models). The order in which the instruments were administered was based on the principle of allowing the participants to respond to the most closed-ended test instrument first (i.e., the PAJT) and the most open-ended test (i.e., the ODCT) at the end. The inclusion of multiple measures was designed to allow the researcher to triangulate the results that would emerge so that a more accurate picture of the participants’ L2 pragmatic development from pretest to posttest could be obtained (Kasper & Roever, 2005).

**Treatment**

The daily schedule of tasks associated with the study treatment is reviewed next. The basic procedure was modeled after that of Izumi (2002). On Day 2, Day 3, and Day 4 of the study, both groups (+ Output and – Output) first viewed the pragmatic video vignettes. Next, learners in both groups were given the opportunity to read transcripts of the video vignettes for approximately one minute. The allowance for reading the transcripts is consistent with Izumi’s (2002) procedure, in which the materials were all texts. Also, having the participants read the transcripts in addition to viewing the vignettes promoted the validity of the treatment task in the sense that what was being targeted was the participants’ exposure to the pragmatic speech acts in conversations. If the participants had only been given the opportunity to view the vignettes, it is possible that the task would have been a test of listening comprehension. Both video/audio and written modes of exposure to the text were included in order to facilitate the comprehension of the materials by all participants. That is, the target of the investigation was not comprehension, so multiple modes of exposure to the input were utilized to promote the study’s focus on the potential effects of Output on the learners’ L2 pragmatic development. A limited amount of
exposure time (i.e., one minute) to the written transcripts was given so that the participants could not simply copy the transcripts and thus rely solely on their written notes, avoiding the need to focus on the key pragmatic elements.

Next, following Izumi (2002), the transcripts of the vignettes were taken away from all participants, so that both the + Output and – Output groups had an identical amount of exposure to the pragmatic materials. At that point, the + Output group received opportunity for output in the form of a text reconstruction task. For half of the reconstruction tasks, each participant, based on her memory of the viewed vignettes and transcripts, worked to reconstruct the dialogue in writing as accurately and completely as possible. The learners were specifically instructed to take notes during the viewing of the videos and time with the transcript, having been advised that they would not have sufficient time to copy down the words verbatim. After the transcripts had been taken away, they were told to write out the exact text of the dialog on the video/transcript with as much accuracy as possible (see Figure 7 below).

The oral reconstruction tasks (the other half of the reconstruction tasks) were similar to the written reconstructions described above, except that the participants were instructed to reconstruct the dialogues from the vignettes orally, by speaking into an audio recorder. The justification for providing opportunities for both written and oral reconstruction stems from the fact that the participants were assessed in both the written and oral modes on the pretest and posttest WDCT and ODCT instruments.

The – Output group did not receive the opportunity for output, but rather answered comprehension-type questions about the video vignettes. The task of answering comprehension questions is analogous to the “extension questions” answered by the – Output groups in Izumi’s (2002) study. A possible concern is that incorporating comprehension questions might, in effect, mean introducing a new variable into the design, but in the case of this study, an actual control group with no treatment at all was ruled out because of the fact that the present study is designed to test the effectiveness of Output in the context of actual instruction. Therefore, the – Output group, rather than doing nothing after viewing the vignettes (an unnatural situation in an ESL classroom), was assigned a typical L2 instructional activity consistent with previous research designs (e.g., Izumi, 2002). The treatment task instructions follow in Figure 7.
On Day 2, both groups viewed the first pragmatic vignette (request refusal). Then, the participants were provided with a copy of the transcript of the vignette to read for one minute. Following their exposure to the transcript, the participants in the + Output group were instructed to reconstruct the text in writing, whereas the – Output group answered comprehension questions about the vignette (see Figure 4). The process was repeated for three more vignettes—an advice-giving vignette, a compliment response, and an invitation vignette. The first and third vignettes (request refusal and compliment response, which are expressive speech acts) were reconstructed in writing, as mentioned above. The second and fourth vignettes (advice-giving and invitation, which are directive speech acts) were reconstructed orally by the learners in the + Output group by speaking into an audio recording device.

On Day 3, the process of viewing and treatment tasks was repeated, with the exception that on this day, the + Output group reconstructed the request refusal and compliment response vignettes orally and reconstructed the other two vignettes (advice-giving and invitation) in writing. Day 4 proceeded in a similar fashion to Day 2, with the same speech acts reconstructed in writing (i.e., request refusal and compliment response) and orally (i.e., advice-giving and invitation). Finally, on Day 5 of the study, four final vignettes were viewed, with the first and third vignette texts reconstructed in writing by the + Output group and the second and fourth...
reconstructed orally. As on the other days, the learners in the – Output group answered comprehension questions after each vignette. In this way, all four speech acts (of both types, expressives and directives) were treated equally and both modes of reconstruction practiced by the + Output group prior to posttesting. After the final treatment had been carried out, the posttest PAJT, WDCT, and ODCT instruments were administered immediately to the + Output and – Output groups.

The choice to use four days of treatment (four treatment sessions of about one hour each) was motivated by the need to balance adequate time for the learners to develop some internalization of the speech acts with efficient treatment and testing that maximized the full participation of the learners. In the existing L2 pragmatic research literature, there is a broad range of treatment time spans. For example, in a study that is often cited as evidence for the effects of instruction on L2 pragmatic development, Olshtain and Cohen (1990) made use of three classroom sessions of only twenty minutes each for the treatment. Takahashi (2005b) implemented a treatment of four ninety-minute sessions. Izumi’s (2002) treatment was six sessions over two weeks. Izumi’s study, however, featured a more complex research design with more variables than in the present study. Other researchers (e.g., Tateyama, 2001) have employed treatment time spans a bit less than these two lengths, and still others (e.g., Ellis, 1992) have conducted longitudinal observations over the course of several months or longer.

In the case of the present study, four treatment sessions of about one hour each were determined by the researcher to be sufficient for the learners to have the opportunity to internalize the speech acts covered in the treatment while avoiding the potential of task fatigue because of the relatively intensive nature of the study tasks (the text reconstruction tasks undertaken by the + Output group and the comprehension and extension questions answered by the participants in the – Output group). Table 14 summarizes the daily instructional treatment and data collection schedule for the two treatment groups. In addition, in order to facilitate a more complete understanding of the treatment procedure, examples of the text reconstructions carried out by two different learners in the + Output group are provided in the next chapter (see Table 17).
Table 14
Daily Schedule for Instructional Treatment and Data Collection

<table>
<thead>
<tr>
<th>Group</th>
<th>Day 1</th>
<th>Day 2 - Day 4</th>
<th>Day 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>• Informed consent</td>
<td>• Video viewing</td>
<td>• Video viewing</td>
</tr>
<tr>
<td></td>
<td>• PAJT pretest</td>
<td>• Text reconstruction (written, oral)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• WDCT pretest</td>
<td></td>
<td>• Posttest PAJT</td>
</tr>
<tr>
<td></td>
<td>• ODCT pretest</td>
<td></td>
<td>• Posttest WDCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Posttest ODCT</td>
</tr>
<tr>
<td>- Output</td>
<td>• Informed consent</td>
<td>• Video viewing</td>
<td>• Video viewing</td>
</tr>
<tr>
<td></td>
<td>• PAJT pretest</td>
<td>• Comprehension questions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• WDCT pretest</td>
<td></td>
<td>• Posttest PAJT</td>
</tr>
<tr>
<td></td>
<td>• ODCT pretest</td>
<td></td>
<td>• Posttest WDCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Posttest ODCT</td>
</tr>
</tbody>
</table>

As mentioned earlier, the data collection and treatment procedure was repeated in a second round roughly four months after the initial round. The number of learners who participated in the first round of data collection and treatment was 17 (7 advanced-level learners in the + Output group, 10 intermediate-level learners in the – Output group). In the second round of data collection, another 17 learners participated (10 intermediate-level learners in the + Output group, 7 advanced-level learners in the – Output group this time).

Scoring and Analyses

Scoring of the Pragmatic Acceptability Judgment Task

The pragmatic acceptability judgment task, following Garcia (2004), is based on individual speech act recognition items. For each item, a situation is presented to the participant,
generally in the form of a dialogue, followed by a series of multiple-choice questions. In Garcia’s study, each item featured four choices. In each case, the learner was expected to select the correct speech act being carried out in the example (prompt). Figure 8 below is an example of a speaker intention-focused PAJT item, taken from Garcia (2004, p. 115):

The woman says, ‘But you’re going to need to check with them and make sure that they remove the rest of that.’ What is the woman doing?

(a) correcting  
(b) requesting  
(c) suggesting  
(d) offering

*Figure 8. Sample speaker intention-focused PAJT item.*

As mentioned above, the PAJT in the present study differed from Garcia’s in that the four choices were actual pragmatic forms (examples of speech acts). In keeping with the research questions, the PAJT items were organized around the four speech acts: request refusals and compliment responses (expressive speech acts) and advice giving and invitations (directives). The participants were instructed to choose the one form that was most appropriate, given the situation. A sample PAJT item from the present study is given in Figure 9 below. The item begins with a brief dialog between a university student and a professor, followed by an opportunity for the participant to choose which of the four options would be the most appropriate response for the student to give. As indicated in the example, the PAJT responses are evaluated as either correct or incorrect. The implications of the PAJT format for the scoring and analysis of these responses is discussed further in the next chapter.
**Student:** Hi, Dr. Smith. You wanted to see me about my paper?

**Professor:** Yes. Your paper is well written. I was wondering if you would be willing to share what you have written with the class.

**Student:**

a) Oh, I’d like to, but I don’t want to do that.
b) Oh, I’d like to, but I won’t because I don’t feel prepared.
c) Oh, I’d like to, but I really don’t feel prepared.
d) Oh, I’d like to, but since I don’t feel prepared, I can’t.

Figure 9. Sample speech act-focused PAJT item.

Individual participants’ scores on the PAJT pretest and posttest instruments were obtained by calculating correct responses for all the items, which reflected the four speech acts tested in the present study (request refusals, advice giving, compliment responses, and invitations). No outside raters were involved in the calculation of the PAJT scores. For the two DCT instruments, however, it was essential to employ outside raters.

**Selection of Raters**

Weir (2005) reported that raters who make use of multi-level criteria, such as the DCT rating scale employed in the present study, are limited in the number of levels that they can practically and consistently distinguish. In order to make use of global scales effectively, raters must be very familiar with the criteria used to assess the responses and, particularly for developing L2 pragmatic research, be able to apply their instincts as native speakers of the L2 to the rating process. Cohen (2004) also discussed the importance of recognizing individual rater characteristics as these influence the raters’ reactions to participant responses.

In order to encourage greater reliability of rating scores, the 28 raters recruited for the present research effort shared three characteristics: (a) all were of similar age to the majority of the study participants (i.e., traditional university age of 18-25 years old); (b) all were native speakers of American English; and (c) all were laypersons with respect to ESL or other L2 instructional experience. These characteristics were in place to encourage ratings of the participant responses that reflected present-day pragmalinguistic norms of North American
English as used in the university context. Utilizing non-experts as raters (unlike what was done in the pilot study) theoretically helped to avoid the problem of the raters’ having too much sympathy for nonnative-like production because of extensive experience working with nonnative speakers of English (McNamara, 1996).

**Rater Training**

The 28 raters were given training in two 30-minute sessions over two weeks. The raters were exposed to examples of both written and oral DCT responses, and practice in rating those responses was carried out in pairs and as a large group. The most important part of the rater training that the raters underwent was familiarization with the four-point rating scale and example responses that correspond to each level (see Tables 12 and 13). In this way, the raters were given the opportunity to become very familiar with the scale and to rate several sample responses from the pilot study data and discuss any discrepancies with fellow raters and the researcher. In order to further facilitate the rater training, a training guide was produced and utilized during the rater training session (see Appendix G).

**Scoring Scale for the DCT Tasks**

The scoring system developed for scoring the DCT responses is analogous to the Next Generation/internet-Based TOEFL 0-4 scale offered by Educational Testing Services (ETS, 2004), with the exception that the 0-level has been eliminated in order to eliminate a “middle rating,” which is theorized to lead to occasional default ratings (e.g., on a 0-4 scale, the score of 2, as the middle score on the scale, may be selected by raters when it is difficult to judge a response). Kasper and Roever (2005) have endorsed rating scales for use in pragmalinguistic assessment, noting that Hudson, Detmer, and Brown (1995), among others, used rating scales profitably to allow external raters to evaluate the pragmatic performance of learners in a language testing context. Also, Cohen (2004) noted that self-assessment rating scales have been developed and used to good effect by researchers such as Brown (2001). Additionally, Yamashita (1996) made use of scoring scales, some of which were adapted from Hudson, Detmer, and Brown (1995) and others, to evaluate the interlanguage pragmatics of American university-age learners of Japanese as a foreign language. In Yamashita’s (1996) study of JFL pragmatics, DCT responses were judged on a 1-5 scale, with the general descriptors of “very unsatisfactory” and “completely appropriate” at either end of the scale. Generally, rating scales for assessing DCT responses have tended to be
impressionistic and somewhat vague. Therefore, for the present study, an important goal is the establishment of more explicit descriptors or indicators for each of the levels of the scale. In the pilot study, the same four-level scale was used by two expert raters with an acceptable level of interrater reliability—agreement on 94.1% of DCT response ratings that were graded by both raters, with a standardized item alpha of .81 (acceptable reliability in social science research; Bachman, 2004; Weir, 2005).

The four-point scale adopted for the present study was motivated by the need to have a holistic scale with which non-expert raters could be familiarized through training in order to produce consistent, reliable scoring of learner responses. Weir (2005) noted that a good rubric should have several key characteristics: (a) clarity in terms of the requirements of the learners; (b) short, simple descriptors; (c) high level of familiarity; (d) specific, accurate instructions. For the purposes of the present study, the characteristics of familiarity and clarity were honed in the rater training session. The rater training guide (Appendix G) provided for hands-on practice that enabled the non-expert raters to identify with the levels of the rating scale.

The other characteristics cited by Weir (2005) appear to be more dependent on the design of the rubric itself. The use of short, simple descriptors and specific, accurate instructions was a central goal in the design the rubric utilized in the present study. Promotion of the internal consistency of a given rating rubric (i.e., the use of like categories of descriptors across levels) is one important measure that research designers need to take in order to facilitate the easy, accessible use of the rubric. Because of the importance of the internal consistency of the rating scale to a valid measurement, a closer look at the rubric proposed for the study is in order at this point. The four levels used for rating the written and oral DCTs in the present study are presented in Table 15, along with a general description of each level and indicators that are intended to help flesh out the characteristics of each level. Familiarity with these descriptions and indicators was the goal of the rater training program mentioned above. A notable contrast in the scoring levels occurs between the scores of 2 and 3. Whereas a score of 2 is generally unacceptable, a score of 3 is generally appropriate. Thus, a score of 3 or 4 is indicative of a generally acceptable or completely acceptable response, while scores of 1 or 2 indicate either generally unacceptable or unacceptable pragmatic responses.
<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>response is completely acceptable pragmatically given the context, not noticeably affected by any errors</td>
<td>- approaches native-like usage - minor grammatical errors do not interfere with pragmatic effectiveness - totally appropriate to the situation</td>
</tr>
<tr>
<td>3</td>
<td>response is generally appropriate given the context, but contains one or more pragmalinguistic flaws that affect the intended meaning</td>
<td>- near native-like usage - minor grammatical errors may distract from pragmatic effectiveness - may be too brief or too long - somewhat appropriate to the situation</td>
</tr>
<tr>
<td>2</td>
<td>response is generally unacceptable pragmatically in this context, though perhaps not in all contexts</td>
<td>- generally non-native-like usage - noticeable errors distract from pragmatic effectiveness at times - too brief or too long - generally inappropriate to the situation</td>
</tr>
<tr>
<td>1</td>
<td>response is unacceptable pragmatically given the context</td>
<td>- clearly non-native-like usage - numerous errors distract from pragmatic effectiveness throughout - clearly inappropriate to the situation</td>
</tr>
</tbody>
</table>

The four-level system observed in the rubric was adopted because of the belief that the relatively simple design of the scale would make the subjective nature of the scoring less problematic. Also, the small number of levels was believed to be conducive to the raters’ grasping the intuitive scale. Finally, as mentioned above, an advantage of a four-level scale for rating responses is that it avoids the common problem of the over-represented middle choice, which might indicate a possible overlapping or item discrimination problem in the response options (Bachman, 2004; Creswell, 2005).
It should be noted that the descriptions and indicators for each level of the scale were
designed to be comparable across those levels. Thus, the indicators at each level cover the areas
of native-like usage, grammatical errors as they affect pragmatic effectiveness, and
appropriateness to the situation within which the learner’s response was given. In addition, the
indicators for levels 2 and 3 require the rater to consider the length of the utterance. This
distinction is designed to help the raters distinguish between what could otherwise be similar
responses. Similarly, the descriptions at each level involve the concepts of acceptability and the
effects of any errors that may be present in the learner’s response. Further clarification of the
four points on the scale can be obtained from a consideration of example utterances typical of
each level (actual responses taken from the pilot study data) in Table 16.

<table>
<thead>
<tr>
<th>Score</th>
<th>Example Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Oh man! I am really sorry because I wasn't in that class, too. Maybe you need to ask somebody else. [request refusal; borrowing class notes]</td>
</tr>
<tr>
<td>3</td>
<td>Thank you, man! To do these, it took 8 hours. [compliment response; new hairstyle]</td>
</tr>
<tr>
<td>2</td>
<td>Thank you. I payed [sic] a lot of money for my hairstyle. What hairstyle do you like? [compliment response; new hairstyle]</td>
</tr>
<tr>
<td>1</td>
<td>Thank you to invite me. The food was really good. Could you invite me next time? [compliment response; good cook]</td>
</tr>
</tbody>
</table>

Having been familiarized with the rating scale and given opportunities to use the scale in
a trial environment during rater training, the raters were charged with rating the WDCT and
ODCT responses. The researcher distributed the WDCT and ODCT responses among the raters,
and each response was scored, item-by-item, by two raters. Because there were 132 DCTs in
total (34 pretest WDCTs, 32 pretest ODCTs, 34 posttest WDCTs, and 32 posttest ODCTs), each
rater scored 9 or 10 DCTs over a three-week period of time (recalling that each DCT must be scored twice, each time by a different rater). The ODCT responses were contained on audio compact discs, and the raters were provided with a scoring sheet for those responses (see Appendix H). The maximum score for any given twelve-item DCT was 48, or an average item score of 4.0. When all ratings had been completed and recorded, the WDCT and ODCT scores were available for analysis and interpretation.

**Statistical Analyses**

Following the scoring of all the data from the three outcome measures, statistical analyses were conducted in order to facilitate accurate interpretation of the data. In order to address Research Question 1, repeated measures ANOVA analyses were carried out in order to determine whether significant increases from pretest to posttest, indicated by significant effects or interactions, were observed on the PAJT, WDCT, and ODCT measures (Bachman, 2004; Creswell, 2005). The ANOVA procedure was selected because of the robustness of this test in determining whether or not the + Output and – Output groups differed (Mackey & Gass, 2005). Following the ANOVAs, individual paired samples $t$-tests were conducted in order to explore and identify the source of any significant effects or interactions (O’Sullivan, 2002). The need for follow-up $t$-tests stems from the fact that the statistical software package used for the analysis does not conduct post-hoc analyses (e.g., Tukey’s HSD) for two treatment group designs such as that in the present study. Finally, effect sizes were calculated for both groups on each instrument (pretest to posttest) because of the tendency of tests of significance that compare means (e.g., the ANOVA and $t$-test) to be tied to sample size (Cohen, 1988).

In order to address the second research question, individual repeated measures ANOVA analyses were carried out (Bachman, 2004; Creswell, 2005). First, all items from the WDCT and ODCT that corresponded to the directive speech acts (i.e., advice-giving and invitations) were separated from those corresponding to the expressive speech acts. Next, all of the + Output ODCT and WDCT responses were subjected to the repeated measures ANOVA analyses to reveal whether or not the output-based instruction and treatment affect the two categories of speech act tested differently. Follow-up $t$-tests were conducted as needed to identify the source of any significant effects or interactions revealed by the ANOVA tests.

As with the analysis of data for the first research question, effect sizes for each instrument on the directive and expressive speech act items were calculated in order to address
the second research question as well. It should be noted that the use of t-tests for identifying the source of significant effects revealed by ANOVA and the use of effect size measures, particularly with respect to small sample sizes, is supported by statistical theory as it applies to social science research (A. Kamata, personal communication, January 24, 2007).

The third research question was addressed by calculating Cronbach’s alpha ($\alpha$) for each set of rater’s scores on each of the DCT instruments; i.e., the scores assigned by two raters for each item on the pretest and posttest WDCT and ODCT responses (Bachman, 2004; McNamara, 1996; Weir, 2005). Such a measure of inter-rater reliability allowed the researcher to determine whether or not the rating scale and the training provided for the raters enabled them to rate the items in a consistent, reliable manner.

**Interpretation of DCT Scores**

Care must be taken in interpreting the results of the written and oral DCTs. Bardovi-Harlig and Hartford (2005) have pointed out that pre-designed research (i.e., research that looks at elicited data rather than naturally occurring language) “may set up a dichotomy between native and non-native speakers” (p. 2). In order to avoid confusing pragmatic acceptability with simply comparing the ESL learners to a vague “native-speaker standard,” the present study featured carefully constructed indicators for use in rating the learner responses on the production-oriented instruments (the written and oral DCTs; see Table 10). However, in the area of pragmatic judgments, it is not possible to entirely eliminate an element of subjectivity because of the fact that a large part of a response’s pragmatic acceptability is its appropriateness in a given social and cultural context. For that reason, it was important that raters in the present study be trained as to the nature of the rating scale and instructed to recognize their own sense of pragmatic acceptability in English. Details concerning rater selection and the rater training process are offered in the next chapter.

**Summary**

The third chapter has presented the major elements of research design and methodology for the present study. The chapter began with a presentation of the pilot study and its limitations and possibilities, including consideration of how the pilot study contributed to the understanding needed to design and carry out the principal study successfully. Next, in a detailed discussion of
the research design of the present study, sections reviewing the participants, materials, procedure and scoring and analyses used in the study were offered. The statistical analyses used to address each of the research questions were explored in some detail. Finally, the interpretation of the DCT rating scale and the selection and training of raters was reviewed.
CHAPTER 4

RESULTS

In the first three chapters of this study, a case has been made for the exploration of an effect for video-based instruction that includes opportunities for output on the development of second language pragmatic competence, based on the theoretical framework of Swain’s (1985, 1995, 2005) Output Hypothesis, applied to the investigation of acquisitional pragmatics (Bardovi-Harlig, 1999). Chapter 3 detailed the setting and participants of the study, outlined the procedure for the treatment and the collection of data and provided the techniques for rating the oral and written DCTs and analyzing the data for each of the pretest-posttest instruments. This chapter presents the results of the study, addressing the research questions raised in Chapter 1 by examining the data for each task in the study as well as the reliability of the DCT rating rubric. The issues of effects and interactions, statistical significance, reliability, and internal consistency of ratings are discussed as the data are presented for each task (PAJT, WDCT, ODCT) and the rating rubric.

Research Questions Revisited

By way of review, each of the research questions is presented below, followed by a discussion of the statistical tests employed to address each questions and the presentation of the pretest-posttest results, organized by task.

1. Is there an effect for Output in the context of video-based pragmatic instruction (including output-focused tasks) on the developing L2 pragmatic competence of adult ESL learners, as evidenced by the acquisition of pragmalinguistic forms (pragmatic speech acts) in appropriate contexts?
   1a. Is there an effect for Output on pragmatic perception on a pragmatic acceptability judgment task (PAJT; Garcia, 2004) from pretest to posttest?
   1b. Is there an effect for Output on written pragmalinguistic production on a written discourse completion task (WDCT) from pretest to posttest?
   1c. Is there an effect for Output on oral pragmalinguistic production on an oral
discourse completion task (ODCT) from pretest to posttest?

2. Is there a differential effect of Output on the two different types of pragmatic speech acts of directives and expressives?
   
   2a. Is there a differential effect for Output on the perception of directives and expressives on a pragmatic acceptability judgment task (PAJT) pretest and posttest?

   2b. Is there a differential effect for Output on the production of directives and expressives on a written discourse completion task (WDCT) pretest and posttest?

   2c. Is there a differential effect for Output on the production of directives and expressives on an oral discourse completion task (ODCT) pretest and posttest?

3. Is the proposed rubric a valid and reliable instrument for the assessment of the pragmatic acceptability of the linguistic production of adult ESL learners on discourse completion tasks (DCTs)?

Text Reconstruction

As discussed in the previous chapter, the + Output group engaged in oral and written text reconstruction tasks following their exposure to the input (video vignettes and transcripts). In order to present an accurate picture of the reconstruction process, Table 17 presents two typical text reconstructions, taken from the actual notes of the participants. One of the reconstructions in the table below was produced by an advanced-level learner and is a reconstruction of a directive speech act. The second reconstruction presented was produced by an intermediate-level learner and is a reconstruction of an expressive speech act. It is notable that in both cases, the learners tended to reconstruct the text with greater accuracy at the beginning of the text, while portions of the text occurring later tended to be summarized. The importance of attention to the pragmalinguistic forms within the texts is highlighted by the fact that the second reconstruction, produced by an intermediate-level learner later in the week, revealed more accuracy of expression of the targeted speech act than the first example, which was produced by an advanced-level learner earlier in the week of treatment.
## Statistical Tests

As was mentioned in the previous chapter, in addition to descriptive statistics, three statistical tests were conducted on the responses and ratings. First, repeated measures ANOVA tests were conducted for all of the participants’ responses on the PAJT, WDCT, and ODCT. Next, follow-up $t$-tests were conducted when necessary to identify the source of any observed effects or interactions indicated in the ANOVA. In addition, effect sizes, expressed in terms of Cohen’s $d$, were calculated for all effects on the three instruments (pretest-posttest). Finally, the

---

### Table 17

**Text Reconstruction Examples**

<table>
<thead>
<tr>
<th>Speech Act</th>
<th>Reconstruction</th>
</tr>
</thead>
</table>
| Advanced learner; directive (advice) | -Hi, how can I help you, sir?  
-Yeah, I want to open an account but I'm not sure what type could you help me out?  
-Yeah, let me ask you a few questions to see what kind of account you need.  
-Okay  
-Would you like to make withdraws anytime?  
-Not really, just probably leave it the money in most of the year.  
-So, would you like maybe three times a year?  
-Not really, I want to withdraw the money every time I want.  
-OK. If I was you I suggest get market account  
-Isn’t that riskier than a regular saving account?  
-No  
-Alright, great, lets go with that then  
-OK |
| Intermediate learner; expressive (request refusal) | -OK everyone, would you please open your books to page 37 please.  
-Let’s do questions on page 37  
-Dr. K  
-Yes, D.  
-I was wondering can we review the question  
-That would be a good idea but I gave everyone 2 days to finish that assignment  
-I’m sorry I forgot what you said  
-Now D., open your book page 37 |
reliability of the rating rubric was tested by calculating Cronbach’s alpha for the two ratings on each item response for both DCT instruments. Each of the statistical tests conducted is explained in more detail below, and a guide to interpreting the results of each test in the context of the present study is presented.

**Repeated Measures ANOVA**

The comparison of more than two groups’ mean performances requires the implementation of a statistical test such as the Analysis of Variance (ANOVA) to determine the ratio of variance within the groups to variance between the groups (the $F$ statistic). Although there are two treatment groups in the present study (+ Output and – Output), it must be remembered that each group actually represents two different conditions: one before the treatment, and one after the treatment. Thus, the repeated measures ANOVA is a more robust measure for identifying significantly different pretest-posttest performance for each group than a simple two-group $t$-test would be (Mackey & Gass, 2005, pp. 274, 277). The repeated measures ANOVA and follow-up $t$-tests described below were conducted using the Statistical Package for the Social Sciences (SPSS), version 11.5, and alpha levels for all tests were set at 0.5, as is customary in second language research (Mackey & Gass, p. 267).

**Follow-up $t$-tests**

Repeated measures ANOVA provides useful information regarding the differences in means as well as any significant effects for the treatment and interactions among the within-group and between-group variables. However, the ANOVA, when carried out utilizing SPSS, does not allow researchers to perform post-hoc calculations (e.g., Tukey’s HSD) on two treatment groups, as is the case in this study. As a result, when the repeated measures ANOVA revealed a significant effect or interaction, follow-up paired samples $t$-tests were conducted to identify the specific source of the effect or interaction indicated. This procedure has precedent in the psycholinguistic research literature (e.g., Rhodes & Anastasi, 2000) and in the second language assessment research literature (e.g., O’Sullivan, 2002).

**Effect Size Calculations**

As Cohen (1988) noted, the need for effect size calculations is important given the tendency of methods of mean comparison (e.g., ANOVA and $t$-tests) to be tied to a researcher’s original measurement unit. Effect size, expressed as Cohen’s $d$, provides a number that indicates the magnitude of an effect (pretest-posttest) in terms of standard deviation. Because the effect
size calculation uses the original scale units in the numerator and denominator, the resulting number is independent of the original scale (pp. 20-21). In the case of the present study, calculating the effect size is particularly important because of the small maximum item score on the rating scale (i.e., 4) and because of the relatively small sample group size (n = 17 per treatment group). Larger sample sizes tend to lead to statistical significance when comparing means. The effect size statistic provides another measure to aid the researcher in assessing the effect of a treatment (American Psychological Association, 2001, p. 25; A. Kamata, personal communication, January 24, 2007).

Cohen (1988) has addressed the interpretation of effect size calculations as well. Noting that the significance of an effect size calculation result is dependent on the particular branch of behavioral science, Cohen offered some general guidelines for understanding the relative size of given effect size statistic. He suggested that a $d$ statistic of .2 or less could be interpreted as a small effect size, although he noted that effect sizes tend to appear smaller in new areas of research in which scales are not well established (p. 25). It could certainly be argued that because the pragmatic rating rubric used in the present study is relatively untested (aside from the pilot study), large effect sizes may be difficult to come by for the DCT instrument ratings. Cohen suggested that a $d$ statistic of .5 might be interpreted as a medium effect size. He added that a medium effect size is one that is likely to be visible upon visual inspection of results. Finally, he suggested that a $d$ statistic of .8 or greater could be interpreted as a large effect size (p. 26).

**Cronbach’s Alpha**

The internal consistency of ratings when more than one rater is responsible to evaluate responses is a key part of the reliability and validity of a rating rubric. The coefficient alpha, also called Cronbach’s alpha, is a good test of internal consistency when the rating scale allows raters to choose from among continuous categories, as is the case with the present study (Creswell, 2005; Cronbach, 1984; Mackey & Gass, 2005). The coefficient alpha was calculated for all pairs of ratings for the WDCT and ODCT pretests and posttests. In keeping with acceptable social science research practices, an alpha of .7 or larger would be regarded as representing sufficiently high inter-rater reliability, as it would indicate an acceptable level of internal consistency among the raters’ scores (Creswell, 2005; Vogt, 2005, p. 71).
The Pragmatic Acceptability Judgment Task

The Pragmatic Acceptability Judgment Task was designed to measure the ability of the participants to recognize which speech act form among four choices was the most appropriate option in response to a given prompt. Research Question 1a asked about the possible effects of output in the context of the video-based instructional treatment on the participants’ pragmatic perception, and was addressed by examining the pretest-posttest data for each group on the judgment task (PAJT). Before examining the results of the repeated measures ANOVA, follow-up *t*-tests and effect size calculations, it is important to consider the descriptive statistics in order to identify general trends in the PAJT pretest-posttest data. For instruments that produce relatively straightforward numerical scores, reporting the mean for each administration can provide a picture of the learners’ performance that is intuitively intelligible (Marsh, 1988). In particular, consideration of any changes in the mean scores for the two Output groups from pretest to posttest may reveal tendencies in the data that can be explored more deeply in the statistical tests. The descriptive statistics for the PAJT are described in Table 18 below. Interestingly, both groups had identical means on the pretest but the + Output group posttest mean was greater than that of the – Output group by a margin of 0.647.

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>7.353</td>
<td>2.473</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>7.353</td>
<td>2.206</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>7.353</td>
<td>2.308</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>8.588</td>
<td>1.544</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>7.941</td>
<td>2.249</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>8.265</td>
<td>1.928</td>
<td>34</td>
</tr>
</tbody>
</table>

*Note. Maximum score on each administration was 12.*
The first test conducted on the PAJT data was a repeated measures ANOVA. The ANOVA revealed a significant effect for Time, $F(1, 32) = 5.284, p < .05$, but no significant Output × Time interaction and no effect for Output (group). As explained in the description of the analytical approach above, this significant time effect would be explored further by conducting a follow-up $t$-test to determine the source of that effect (see Table 20). Table 19 provides the details of the ANOVA results for the PAJT response scores for each of the two Output groups.

Table 19
**PAJT ANOVA Table**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>1.779</td>
<td>1</td>
<td>1.779</td>
<td>.272</td>
</tr>
<tr>
<td>Time</td>
<td>14.132</td>
<td>1</td>
<td>14.132</td>
<td>5.284*</td>
</tr>
<tr>
<td>Output × Time</td>
<td>1.779</td>
<td>1</td>
<td>1.779</td>
<td>.665</td>
</tr>
<tr>
<td>Error</td>
<td>85.588</td>
<td>32</td>
<td>2.675</td>
<td></td>
</tr>
</tbody>
</table>

* $p < .05$

A follow-up paired-samples $t$-test revealed that the source of the Time effect was the pretest-posttest change among the + Output group, $t(16) = 2.452, p < .05$, two-tailed. As indicated in Table 20, the small $p$-value observed for the + Output group pretest-posttest (.026) stood in sharp contrast to the large $p$-value observed for the – Output group (.351).

Table 20
**PAJT t-test Table**

<table>
<thead>
<tr>
<th>Source</th>
<th>$t$</th>
<th>df</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>2.452</td>
<td>16</td>
<td>.026</td>
</tr>
<tr>
<td>- Output</td>
<td>.960</td>
<td>16</td>
<td>.351</td>
</tr>
</tbody>
</table>
The effect size for Output was also calculated. As mentioned at the beginning of this chapter, the inclusion of the effect size statistic is important in the present study because of the relatively limited sample size ($n = 17$ per treatment group). As has been noted (e.g., Cohen, 1988; Norris & Ortega, 2000), tests for statistical significance such as the repeated measures ANOVA and $t$-test are highly influenced by sample size, with larger sample sizes tending to yield statistically significant test results, all other things being equal. Ellis (2000) noted that the reporting of effect sizes also facilitates the comparison of results and findings across studies, thus enabling second language researchers to draw conclusions as to the relative importance of findings over time. For that reason, effect sizes were consistently calculated for all pretest-posttest measures in the present study.

The effect sizes calculated for the + Output and – Output groups on the PAJT pretest-posttest, expressed in terms of Cohen’s $d$, are presented in Table 21. As indicated in the table, the effect size associated with the + Output treatment was substantial, much greater than the .8 level normally considered to represent a large effect size. The effect size for the – Output group was not nearly as large. In fact, the – Output effect size was below the .5 level normally associated with a medium effect size (Cohen, 1988). The effect size calculations for the overall PAJT results confirmed the findings of the repeated measures ANOVA and follow-up $t$-test that the + Output group performed significantly better on the overall PAJT task than the – Output group from pretest to posttest.

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>1.226</td>
</tr>
<tr>
<td>- Output</td>
<td>.480</td>
</tr>
</tbody>
</table>

Question 2a concerned the possible differential effect of Output on directive and expressive speech acts in perception on the PAJT. This question was addressed by examining the
pretest-posttest data for the directive and expressive speech acts for each of the two treatment groups. The descriptive statistics for the directive speech act items on the PAJT for both groups are detailed in Table 22.

Table 22
PAJT Directives Descriptive Statistics

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>3.000</td>
<td>1.541</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>2.882</td>
<td>1.269</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>2.941</td>
<td>1.391</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>3.882</td>
<td>1.054</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>3.529</td>
<td>1.546</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>3.706</td>
<td>1.315</td>
<td>34</td>
</tr>
</tbody>
</table>

Note. Maximum score on each administration was 6.

Next, each of the tests conducted in addressing the second research question as it related to the PAJT is discussed below. First, the repeated measures ANOVA for the PAJT directives revealed a significant effect for Time, \( F(1, 32) = 7.988, p < .01 \), but no significant Output \( \times \) Time interaction or significant effect for Output. The significant effect for Time was explored further in a follow-up \( t \)-test. Table 23 details the results of the ANOVA.

Table 23
PAJT Directives ANOVA Table

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>.941</td>
<td>1</td>
<td>.941</td>
<td>.377</td>
</tr>
<tr>
<td>Time</td>
<td>9.941</td>
<td>1</td>
<td>9.941</td>
<td>7.988**</td>
</tr>
<tr>
<td>Output ( \times ) Time</td>
<td>.235</td>
<td>1</td>
<td>.235</td>
<td>.189</td>
</tr>
<tr>
<td>Error</td>
<td>39.824</td>
<td>32</td>
<td>1.244</td>
<td></td>
</tr>
</tbody>
</table>

** \( p < .01 \)
A follow-up $t$-test revealed that the source of the significant effect for Time was the pretest-posttest increase among the $+$ Output group, $t(16) = 2.504, p < .05$, two-tailed, as indicated in Table 24.

<table>
<thead>
<tr>
<th>Source</th>
<th>$T$</th>
<th>$df$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+$ Output</td>
<td>2.504</td>
<td>16</td>
<td>.023</td>
</tr>
<tr>
<td>$-$ Output</td>
<td>1.575</td>
<td>16</td>
<td>.135</td>
</tr>
</tbody>
</table>

The calculation of the effect sizes for each group with respect to the directive speech acts revealed a quite large effect size for the $+$ Output group ($d = 1.252$) and a medium effect size for the $-$ Output group ($d = .788$). This finding is consistent with the results of the repeated measures ANOVA and $t$-test measures in supporting superior performance by the $+$ Output group from pretest to posttest. The effect size statistics are displayed in Table 25.

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+$ Output</td>
<td>1.252</td>
</tr>
<tr>
<td>$-$ Output</td>
<td>.788</td>
</tr>
</tbody>
</table>

The responses to the expressive speech act items on the PAJT were also analyzed in the effort to address Research Question 2a. For the six expressive items on the pretest, the $+$ Output group and $-$ Output group differed only by .1177 on average. On the posttest, the $+$ Output group demonstrated a mean gain of .353 per item, whereas the $-$ Output group showed a negligible
decrease (-.0588). This difference in means, along with other descriptive statistics, is described in Table 26.

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>4.353</td>
<td>1.367</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>4.471</td>
<td>1.231</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>4.412</td>
<td>1.282</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>4.706</td>
<td>.849</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>4.412</td>
<td>1.064</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>4.559</td>
<td>.960</td>
<td>34</td>
</tr>
</tbody>
</table>

*Note.* Maximum score on each administration was 6.

The repeated measures ANOVA that tested the pretest-posttest performances of the two treatment groups with respect to the expressive speech acts revealed no significant effects or interactions. Table 27 provides details for the PAJT expressives ANOVA.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>.132</td>
<td>1</td>
<td>.132</td>
<td>.075</td>
</tr>
<tr>
<td>Time</td>
<td>.368</td>
<td>1</td>
<td>.368</td>
<td>.429</td>
</tr>
<tr>
<td>Output × Time</td>
<td>.721</td>
<td>1</td>
<td>.721</td>
<td>.841</td>
</tr>
<tr>
<td>Error</td>
<td>27.412</td>
<td>32</td>
<td>.857</td>
<td></td>
</tr>
</tbody>
</table>

Because no significant effect or interaction was observed, no follow-up t-test was conducted. Effect sizes for both groups were calculated with respect to the expressives, revealing
a medium effect size for the + Output group ($d = .622$) and a markedly smaller effect size for the – Output group ($d = .085$), as displayed in Table 28. As has been mentioned, effect size calculations may reveal not only the magnitude of an effect, but also something about the source of the effect based on the fact that the calculation makes use of the $t$-statistics from the two different treatment groups’ pretest-posttest comparisons (Cohen, 1988). The large difference between the effect sizes for the + Output and – Output groups from pretest to posttest is thus noteworthy.

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>.622</td>
</tr>
<tr>
<td>- Output</td>
<td>.085</td>
</tr>
</tbody>
</table>

The Written Discourse Completion Task

The written DCT data (pretest-posttest) provide insight into Research Question 1b. Recall that the WDCT was designed to determine the extent to which participants could express in writing pragmatically appropriate speech acts in response to written situational prompts. It should be kept in mind that unlike in the case of the PAJT responses, what is being analyzed on the DCT tasks (both the WDCT and the ODCT) are the raters’ assessments of the participants’ responses, all of which were rated on a 4-point scale. An initial look at the overall WDCT data indicated no significant effects or interactions, although exploration of the individual speech act types (directives and expressives) yielded several interesting findings that warranted closer investigation.

The descriptive statistics for the overall WDCT results appear in Table 29. As indicated in the descriptive statistics table, the changes in means from pretest to posttest were relatively small for both treatment groups, although it is noteworthy that the mean change for the + Output group was positive and was clearly greater than the small decrease observed for the – Output
group (an increase of 3.294 as compared with a mean decrease of 0.353). This contrast would be shown to be indicative of some observable differences between the two groups upon closer investigation of the two speech act types. First, however

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>33.000</td>
<td>5.117</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>35.471</td>
<td>5.533</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>34.235</td>
<td>5.395</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>36.294</td>
<td>5.489</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>35.118</td>
<td>4.745</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>35.706</td>
<td>5.087</td>
<td>34</td>
</tr>
</tbody>
</table>

Note. Maximum score on each administration was 48 (12 items, 4 possible points on each).

Next, a repeated measures ANOVA was conducted to determine whether any significant effects or interactions might be observed for the overall WDCT data (see Table 30).

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>7.118</td>
<td>1</td>
<td>7.118</td>
<td>.187</td>
</tr>
<tr>
<td>Time</td>
<td>36.765</td>
<td>1</td>
<td>36.765</td>
<td>2.210</td>
</tr>
<tr>
<td>Output × Time</td>
<td>56.529</td>
<td>1</td>
<td>56.529</td>
<td>3.397</td>
</tr>
<tr>
<td>Error</td>
<td>532.456</td>
<td>32</td>
<td>16.639</td>
<td></td>
</tr>
</tbody>
</table>
The repeated measures ANOVA revealed no statistically significant effects or interactions for the overall WDCT pretest-posttest data, as indicated in the ANOVA table. Because no significant effect or interaction was observed, no follow-up paired-samples t-test was determined to be necessary. Effect sizes for both treatment groups were calculated in order to confirm the trend toward non-significance observed for this task. Table 31 indicates that Cohen’s $d$ calculated for the + Output group was .965, and for the – Output group, Cohen’s $d$ was calculated at .177.

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>.965</td>
</tr>
<tr>
<td>- Output</td>
<td>.177</td>
</tr>
</tbody>
</table>

Surprisingly, the effect size calculations revealed that the effect of the treatment on the + Output group was relatively large ($d = .965$), despite the lack of statistical significance identified by the repeated measures ANOVA. The effect size for the – Output group was quite small ($d = .177$), as was expected. The results of the calculation of Cohen’s $d$ indicated a large effect for the + Output instructional treatment that would be tested next as to its effect on the participants’ pretest-posttest performance on the directive speech act items on the WDCT.

In addressing Research Question 2b and continuing to clarify the possible effect of the + Output treatment for the WDCT, the participants’ responses on the WDCT directive speech act items were examined. Table 32 presents the descriptive statistics for the WDCT directives for the + Output and – Output groups. As the table indicates, the + Output group again demonstrated a clear mean increase from the pretest to the posttest (1.294 on a 24-point instrument). By contrast, the – Output group exhibited a mean decrease of .265.
Table 32

WDCT Directives Descriptive Statistics

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>16.706</td>
<td>3.143</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>17.500</td>
<td>2.704</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>17.103</td>
<td>2.915</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>18.000</td>
<td>2.937</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>17.235</td>
<td>3.255</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>17.618</td>
<td>3.077</td>
<td>34</td>
</tr>
</tbody>
</table>

Note. Maximum score on each administration was 24 (6 items, 4 possible points on each).

In order to determine whether the mean increase observed for the + Output group was indicative of a significant effect for the Output treatment, a repeated measures ANOVA was carried out. The results of the ANOVA test for the WDCT directive responses are presented in Table 33.

Table 33

WDCT Directives ANOVA Table

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>.004</td>
<td>1</td>
<td>.004</td>
<td>.986</td>
</tr>
<tr>
<td>Time</td>
<td>4.504</td>
<td>1</td>
<td>4.504</td>
<td>.737</td>
</tr>
<tr>
<td>Output × Time</td>
<td>10.327</td>
<td>1</td>
<td>10.327</td>
<td>1.690</td>
</tr>
<tr>
<td>Error</td>
<td>195.544</td>
<td>32</td>
<td>6.111</td>
<td></td>
</tr>
</tbody>
</table>

No follow-up t-test was determined to be necessary to identify the source of any effect or interaction, as no statistically significant effect or interaction was observed on the repeated measures ANOVA. The effect sizes for both treatment groups were calculated, however. The results of those calculations are detailed in Table 34.
Table 34

WDCT Directives Effect Sizes

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>.636</td>
</tr>
<tr>
<td>- Output</td>
<td>.209</td>
</tr>
</tbody>
</table>

The effect size calculations revealed that, with respect to directives on the WDCT, the effect size of the treatment for the + Output group was medium in size ($d = .636$), whereas the effect size calculated for the – Output group was relatively small ($d = .209$). Despite the findings of the ANOVA, there appears to have been at least a medium sized effect for the + Output instructional treatment with respect to the performance of those learners on the WDCT directive items from pretest to posttest.

The next step in addressing Question 2b was the consideration of the WDCT responses for the expressive speech act items. The descriptive statistics for the participants’ scores on the WDCT expressives are presented in Table 35. As was the case with the directive speech act items, the learners in the + Output group demonstrated a clear mean increase from pretest to posttest (2 points on a 24-point instrument) and the – Output group again decreased in mean rating scores for the expressives from pretest to posttest (-.088).

Table 35

WDCT Expressives Descriptive Statistics

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>$M$</th>
<th>$SD$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>16.294</td>
<td>2.884</td>
<td>17</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>17.971</td>
<td>3.194</td>
<td>17</td>
</tr>
<tr>
<td>Total pretest</td>
<td>17.132</td>
<td>3.115</td>
<td>34</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>18.294</td>
<td>2.979</td>
<td>17</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>17.882</td>
<td>2.322</td>
<td>17</td>
</tr>
<tr>
<td>Total posttest</td>
<td>18.088</td>
<td>2.639</td>
<td>34</td>
</tr>
</tbody>
</table>

*Note.* Maximum score on each administration was 24 (6 items, 4 possible points on each).
A repeated measures ANOVA was also carried out on the WDCT pretest-posttest data for the expressive speech acts. As with the WDCT directives, this ANOVA also failed to indicate any statistically significant effects or interactions for either group. Table 36 provides the details of the ANOVA results.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>6.798</td>
<td>1</td>
<td>6.798</td>
<td>.622</td>
</tr>
<tr>
<td>Time</td>
<td>15.533</td>
<td>1</td>
<td>15.533</td>
<td>2.839</td>
</tr>
<tr>
<td>Output × Time</td>
<td>18.533</td>
<td>1</td>
<td>18.533</td>
<td>3.388</td>
</tr>
<tr>
<td>Error</td>
<td>175.059</td>
<td>32</td>
<td>5.471</td>
<td></td>
</tr>
</tbody>
</table>

Although no statistically significant effect or interaction was observed in the ANOVA results, it is worth noting that there was an observable trend in the data, as indicated by the mean change among the participants in the + Output group (see Table 35 above). The effect sizes for both Output groups were calculated in order to understand better the strength of this finding (see Table 37).

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>1.125</td>
</tr>
<tr>
<td>- Output</td>
<td>.902</td>
</tr>
</tbody>
</table>
The effect size calculations revealed that the effect of the treatment for the + Output group was quite large ($d = 1.125$), and that the effect size for the – Output group was also relatively large ($d = .902$). These effect size results confirm the presence of a positive effect for the learners in the + Output treatment group on the WDCT expressive items, as was indicated in the overall descriptive statistics for the + Output group on the WDCT expressives.

The Oral Discourse Completion Task

The ODCT data reflect a different number of participants than the PAJT and WDCT results. Specifically, the oral responses of two participants, one representing the + Output group and one from the – Output group, had to be discarded. The reason for this loss of data was technical in nature. As mentioned in the previous chapter, the learners’ oral responses were spoken into recording devices either on personal computers (using Microsoft® Sound Recorder, version 5.1) or on handheld digital voice recorders (Olympus® models WS-200S and DS-2). Unfortunately, during the audio recording process in the second round of data collection, it was discovered that one of the personal computers being utilized by the learners was recording improperly, resulting in an unacceptably high level of distortion when the participants’ responses were played back. Because one participant from each group had used that particular recording station, it was decided to discard the data from both participants, which left the total sample size at 32 for the ODCT instrument, with 16 + Output and 16 – Output group participants.

The specific technological difficulty that was encountered was unexpected, and every effort was taken before the recording process to see that the equipment was functioning properly. However, such problems represent an inherent risk in carrying out research that is technology-based. Given the large number of digital recordings that were entered, downloaded and saved to audio compact discs in the course of the PAJT data collection, the number of problems was limited. The question of how to limit technological difficulties in second language research will be addressed further in the fifth chapter.

Research Question 1c was addressed by considering the pretest-posttest data for the oral DCT. As a reminder, the scores analyzed in this section are the ratings of the participants as submitted by the undergraduate raters, with each item scored by two raters on a scale from 1 to 4. The descriptive statistics for the overall ODCT data are presented in Table 38. As the table
indicates, the means for both groups increased appreciatively from the pretest to posttest. For the + Output group, the mean increase was 3.219 (on this 48-point instrument). It was found that the – Output group also demonstrated a large mean increase from pretest to posttest (4.281). The significance of these mean differences would be explored further in the subsequent statistical tests carried out on the ODCT data.

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>30.719</td>
<td>6.094</td>
<td>16</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>33.750</td>
<td>6.743</td>
<td>16</td>
</tr>
<tr>
<td>Total pretest</td>
<td>32.234</td>
<td>6.507</td>
<td>32</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>33.938</td>
<td>4.722</td>
<td>16</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>38.031</td>
<td>4.613</td>
<td>16</td>
</tr>
<tr>
<td>Total posttest</td>
<td>35.984</td>
<td>5.041</td>
<td>32</td>
</tr>
</tbody>
</table>

*Note. Maximum score on each administration was 48 (12 items, 4 possible points on each).*

A repeated measures ANOVA carried out on the overall ODCT data revealed a significant effect for Output in the overall ODCT pretest-posttest data, $F(1, 30) = 4.586, p < .05$, and a significant effect for Time as well, $F(1, 30) = 11.963, p < .01$, but no significant Output × Time interaction. These findings suggest a difference between the two groups that may have existed on both their pretest and posttest performances. In fact, the descriptive statistics described in Table 38 above support the assertion that the – Output group exhibited a higher mean rating score than the + Output group on both the ODCT pretest (by a margin of 3.031) and on the posttest (by a margin of 4.094). As for the significant effect observed for time, the mean increases for both groups mentioned above account for that finding. Table 39 presents the details of the repeated measures ANOVA test results mentioned above.
Table 39  
**ODCT ANOVA Table**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>203.062</td>
<td>1</td>
<td>203.062</td>
<td>4.586*</td>
</tr>
<tr>
<td>Time</td>
<td>225.00</td>
<td>1</td>
<td>225.000</td>
<td>11.963**</td>
</tr>
<tr>
<td>Output × Time</td>
<td>4.516</td>
<td>1</td>
<td>4.516</td>
<td>.240</td>
</tr>
<tr>
<td>Error</td>
<td>564.234</td>
<td>30</td>
<td>18.808</td>
<td></td>
</tr>
</tbody>
</table>

* *p < .05  ** *p < .01

The follow-up paired-samples *t*-test carried out to investigate the effects identified in the repeated measures ANOVA revealed that the source of the effects was the pretest-posttest change among the – Output group. Table 40 indicates a significant pretest-posttest effect for the – Output group, *t*(15) = 3.665, *p* < .01, two-tailed.

Table 40  
**ODCT t-test Table**

<table>
<thead>
<tr>
<th>Source</th>
<th><em>t</em></th>
<th>df</th>
<th><em>p</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>1.762</td>
<td>15</td>
<td>.098</td>
</tr>
<tr>
<td>- Output</td>
<td>3.665</td>
<td>15</td>
<td>.002</td>
</tr>
</tbody>
</table>

In order to confirm the magnitude of the effect associated with the – Output treatment, effect sizes were calculated for both groups. As Table 41 indicates, Cohen’s *d* for the + Output group was calculated at .910, indicative of a relatively large effect size, and the – Output instructional treatment effect size was quite large for the ODCT, with *d* = 1.893. The large effect sizes for both groups suggest that both instructional treatments produced meaningful effects for the learners from pretest to posttest. The question of whether or not these large effects were observable for the two groups’ pretest-posttest performances on the ODCT directive and
expressive speech act items would be explored next.

Table 41

<table>
<thead>
<tr>
<th>ODCT Effect Sizes</th>
<th>Group</th>
<th>Effect Size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Output</td>
<td>.910</td>
</tr>
<tr>
<td></td>
<td>- Output</td>
<td>1.893</td>
</tr>
</tbody>
</table>

Research Question 2c concerned the possible differential effect of the output-focused video-based treatment on the oral production of directive and expressive speech acts. The descriptive statistics for the ODCT directive response ratings are detailed in Table 42.

Table 42

<table>
<thead>
<tr>
<th>ODCT Directives Descriptive Statistics</th>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ Output pretest</td>
<td>15.406</td>
<td>3.073</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>- Output pretest</td>
<td>16.906</td>
<td>3.738</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Total pretest</td>
<td>16.156</td>
<td>3.451</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>+ Output posttest</td>
<td>16.156</td>
<td>2.833</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>- Output posttest</td>
<td>18.156</td>
<td>2.688</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Total posttest</td>
<td>17.156</td>
<td>2.900</td>
<td>32</td>
</tr>
</tbody>
</table>

*Note.* Maximum score on each administration was 24 (6 items, 4 possible points on each).

The descriptive statistics revealed mean increases from pretest to posttest for both of the treatment groups. For the + Output group, the mean increase was 0.75, and for the – Output group, the increase was 1.25 on the 24-point instrument. In order to explore the possibility of a significant effect for output, a repeated measures ANOVA was carried out on the ODCT pretest-
posttest performances on the directive speech act items by the participants of both groups. The ANOVA revealed no significant effect or interaction, but there was actually a relatively small $p$-value ($p = .061$) for the Output between-subjects effect. Table 43 provides the details of the ODCT directives ANOVA results.

Table 43

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>49.000</td>
<td>1</td>
<td>49.000</td>
<td>3.784</td>
</tr>
<tr>
<td>Time</td>
<td>16.000</td>
<td>1</td>
<td>16.000</td>
<td>2.507</td>
</tr>
<tr>
<td>Output × Time</td>
<td>1.000</td>
<td>1</td>
<td>1.000</td>
<td>.157</td>
</tr>
<tr>
<td>Error</td>
<td>191.500</td>
<td>30</td>
<td>6.383</td>
<td></td>
</tr>
</tbody>
</table>

In order to investigate further the effects of the treatment on the two groups that may not have been evident in the ANOVA above, Cohen’s $d$ was calculated for both groups. The results of the effect size calculations are detailed in Table 44. As the effect size table makes clear, the effect size for the + Output group was relatively small ($d = .367$), whereas the effect size for the - Output instructional treatment group was large ($d = .933$).

Table 44

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s $d$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>.367</td>
</tr>
<tr>
<td>- Output</td>
<td>.933</td>
</tr>
</tbody>
</table>
In order to continue addressing Research Question 2c, the ODCT responses for the expressive speech act items were analyzed. An initial look at the descriptive statistics, summarized in Table 45, reveals that the mean ratings for both the + Output and – Output groups increased appreciably from the pretest to the posttest instrument. The + Output group demonstrated an increase of 2.3438 and the – Output group exhibited an increase of 2.8437 pretest-posttest on this 24-point instrument. The possibility that both groups were manifesting significant treatment effects was explored through the subsequent statistical tests conducted on the ODCT expressive response ratings.

Table 45

<table>
<thead>
<tr>
<th>Group and Time</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output pretest</td>
<td>15.313</td>
<td>3.321</td>
<td>16</td>
</tr>
<tr>
<td>- Output pretest</td>
<td>17.094</td>
<td>3.508</td>
<td>16</td>
</tr>
<tr>
<td>Total pretest</td>
<td>16.203</td>
<td>3.480</td>
<td>32</td>
</tr>
<tr>
<td>+ Output posttest</td>
<td>17.656</td>
<td>2.322</td>
<td>16</td>
</tr>
<tr>
<td>- Output posttest</td>
<td>19.938</td>
<td>2.167</td>
<td>16</td>
</tr>
<tr>
<td>Total posttest</td>
<td>18.797</td>
<td>2.495</td>
<td>32</td>
</tr>
</tbody>
</table>

Note. Maximum score on each administration was 24 (6 items, 4 possible points on each).

A repeated measures ANOVA was carried out to examine the ODCT pretest-posttest data with respect to the expressive speech act items. Table 46 provides the details of the ANOVA results. The ANOVA revealed significant effects for Output group, $F(1, 30) = 5.791, p < .05$, and for Time, $F(1, 30) = 20.264, p < .001$. The finding that both Output and Time demonstrated significant effects is consistent with the fact that both the + Output group and the – Output group experienced noticeable gains in mean score from pretest to posttest for the expressive items on the Oral Discourse Completion Task.
Table 46

**ODCT Expressives ANOVA Table**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>66.016</td>
<td>1</td>
<td>66.016</td>
<td>5.791*</td>
</tr>
<tr>
<td>Time</td>
<td>107.641</td>
<td>1</td>
<td>107.641</td>
<td>20.264**</td>
</tr>
<tr>
<td>Output × Time</td>
<td>1.000</td>
<td>1</td>
<td>1.000</td>
<td>.188</td>
</tr>
<tr>
<td>Error</td>
<td>159.359</td>
<td>30</td>
<td>5.312</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05   ** p < .001

A follow-up $t$-test was conducted to identify the sources of the effects identified on the repeated measures ANOVA (see Table 47). The $t$-test revealed a significant pretest-posttest effect for both the + Output treatment group, $t(15) = 2.471, p < .05$, two-tailed, and the – Output treatment group, $t(15) = 4.344, p < .01$, two-tailed.

Table 47

**ODCT Expressives $t$-test Table**

<table>
<thead>
<tr>
<th>Source</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>2.471</td>
<td>15</td>
<td>.026</td>
</tr>
<tr>
<td>- Output</td>
<td>4.344</td>
<td>15</td>
<td>.001</td>
</tr>
</tbody>
</table>

The follow-up $t$-test revealed contributions to the time and group effects from both the + Output and – Output groups, with a clearly significant effect contributed from the – Output group. In order to further examine the differences between the two groups with respect to the ODCT expressives, the effect sizes for both groups were calculated. Should the effect sizes prove to be large for either or both groups, the effects identified in the statistical tests above would be confirmed. The effect size calculations would further provide insight into whether the – Output group effect was actually larger than that of the + Output group, as suggested by the $t$-test results. Table 48 provides the details of the effect size calculations.
Table 48

**ODCT Expressives Effect Sizes**

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Size (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Output</td>
<td>1.276</td>
</tr>
<tr>
<td>- Output</td>
<td>2.243</td>
</tr>
</tbody>
</table>

The effect size calculations revealed that both groups had large effect sizes, with the effect size for the – Output group being quite large ($d = 2.243$). The large effects observed for both instructional treatments were consistent with the findings of the $t$-test described in Table 47 above.

**The DCT Rating Rubric**

The DCT rating rubric used in the present study included four levels for assessing the pragmatic acceptability of participants’ responses on the WDCT and ODCT pretest and posttest instruments (see Tables 15 and 16). The scale was utilized by 28 trained, non-expert raters over the course of 3 weeks (following two weeks of training) to generate rating scores for every participant’s item responses on all four DCT instruments (see Appendix G). Given that there were 12 items on each DCT instrument, the maximum rating score possible on an entire instrument was 48.

Research Question 3 concerns the internal consistency of the DCT rating rubric, which is in fact a type of reliability. This question was addressed by calculating the coefficient alpha (or Cronbach’s alpha) for each of the four rated instruments. Coefficient alpha calculations provide a good estimate of the consistency of scores that represent continuous variables, such as those on the DCT rating rubric (Creswell, 2005, p. 164). Cronbach’s alpha, also referred to as coefficient alpha, is particularly suitable for assessing the internal consistency of ordinal data, such as that generated by the rating scale utilized for the present study (Mackey & Gass, 2005, p. 130). Reliability analysis, conducted using the Statistical Package for the Social Sciences (SPSS), version 11.5, revealed acceptable reliability levels (greater than 0.7) for three of the four instruments. Table 49 provides the details of the coefficient alpha calculation results.
Table 49  
*Cronbach’s alpha for WDCT and ODCT Instruments*

<table>
<thead>
<tr>
<th></th>
<th>WDCT Pretest</th>
<th>ODCT Pretest</th>
<th>WDCT Posttest</th>
<th>ODCT Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.885</td>
<td>.795</td>
<td>.768</td>
<td>.632</td>
</tr>
</tbody>
</table>

The relatively high coefficient alpha statistic calculated for the WDCT pretest instrument ratings and the moderately high alphas observed for the ODCT pretest and WDCT posttest provide a measure of confidence in the internal consistency of the rating rubric as utilized by the 28 raters in the present study. The somewhat lower coefficient alpha identified for the ODCT posttest instrument may be an aberration, or may be the result of differences characteristic of the participants’ ODCT posttest responses. This issue is explored further in the fifth chapter.

**Summary**

The examination of the results of the PAJT pretest and posttest suggested that there were significant effects for time, originating with the + Output video-based treatment, on the PAJT pretest-posttest overall and for the PAJT directive items. In addition, there were large effect sizes observed for the + Output group on the PAJT pretest-posttest overall and for the PAJT directives pretest-posttest. A review of the WDCT results indicated that there was a significant pretest-posttest effect for the + Output group on the WDCT expressive item ratings. Also, large effect sizes were observed for the + Output group on the overall WDCT instrument pretest-posttest and on the WDCT expressive items pretest-posttest. The ODCT results revealed that there was a significant effect for time, originating with the – Output treatment group, on the overall ODCT instrument. Also, there were effects for time and group, originating with both the + Output and – Output groups, on the ODCT expressive items. In addition, large effect sizes were observed for both the + Output and – Output groups on the overall ODCT instrument pretest-posttest and on the ODCT expressive item ratings pretest-posttest. Finally, large effect sizes were calculated for the – Output group on the ODCT directive items pretest-posttest.
The examination of the internal consistency of the DCT rating rubric as utilized by the raters for this study revealed an acceptable level of inter-rater reliability for three of the four DCT instruments (the WDCT pretest and posttest and the ODCT pretest), whereas the ODCT posttest ratings exhibited internal consistency that was marginally below the generally accepted level for second language research. Table 50 provides a summary of the findings discussed above, arranged according to the research questions of the study rather than by task.

Table 50
Summary of Research Questions and Findings

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Effect of Output on pragmalinguistic perception (PAJT)?</td>
<td>Statistically significant effect for Output; large effect size for the + Output treatment</td>
</tr>
<tr>
<td>1b. Effect of Output on written pragmalinguistic production (WDCT)?</td>
<td>No statistically significant effect for Output; large effect size for the + Output treatment</td>
</tr>
<tr>
<td>1c. Effect of Output on oral pragmalinguistic production (ODCT)?</td>
<td>No statistically significant effect for Output; significant effect for the – Output treatment; large effect sizes for + Output and – Output treatments</td>
</tr>
<tr>
<td>2a. Different effect for Output on directive and expressive speech act perception (PAJT)?</td>
<td>Statistically significant effect and large effect size for Output on directives; no significant effect and only medium effect size for Output on expressives</td>
</tr>
<tr>
<td>2b. Different effect for Output on directive and expressive speech act written production (WDCT)?</td>
<td>No statistically significant effect (ANOVA) for Output on directives or expressives; notable pretest-posttest mean change (in descriptives) and large effect size for Output on expressives</td>
</tr>
<tr>
<td>2c. Different effect for Output on directive and expressive speech act oral production (ODCT)?</td>
<td>No statistically significant effect for Output on directives; significant effects and large effect sizes for both + Output and - Output treatments on expressives</td>
</tr>
<tr>
<td>3. Internal consistency and reliability of the pragmatic rating rubric?</td>
<td>Acceptable for the WDCT pretest, WDCT posttest, and ODCT pretest; slightly less than acceptable levels for the ODCT posttest</td>
</tr>
</tbody>
</table>
All of the findings summarized in this chapter are explored further in Chapter 5. The fifth chapter organizes the findings according to the research questions for the purpose of discussion and interpretation.
CHAPTER 5

DISCUSSION AND CONCLUSIONS

Overview

Schauer (2006) noted the importance of the environment in shaping learners’ pragmatic awareness in a target language. In that study, ESL learners outperformed their EFL counterparts, reinforcing the findings of Bardovi-Harlig and Dörnyei (1998), the study that Schauer’s research replicated and extended, for pragmatic recognition, or perception, and those of House (1996) for pragmatic production. In the case of the present study, which also examined pragmatic perception and production, the context was instructed adult English as a second language, and the setting was an Intensive English program located on the campus of a large public university in the southeastern United States.

The primary purpose of this study was to determine whether or not there was an effect for an output-based instructional treatment that included pragmatic video vignettes on adult ESL learners’ developing L2 pragmatic perception and production. In addition, the study addressed the internal consistency and reliability of a rating rubric used to assess the pragmatic appropriateness of ESL learners’ responses on DCT instruments.

The participants were generally highly motivated adult learners of English with postsecondary education backgrounds in their home countries. The learners represented ten different first language and culture backgrounds. In order of the number of participants, these were: Korean, Spanish, Arabic, Mandarin Chinese, Brazilian Portuguese, Turkish, French, Italian, Japanese, and Vietnamese. As indicated in Chapter 3 above, the + Output and – Output groups each consisted of two intact ESL classes: one advanced-level class and one intermediate-level class. Each instructional treatment group included 17 learners, although data for only 16 participants in each group were used for the ODCT analysis because of technical problems in the data collection phase of the study.

The present chapter offers interpretation for each of the findings described in Chapter 4, and is organized around the three research questions of the study. In addition, the limitations of the study and recommendations for the application of the results of the study to acquisitional
pragmatics research and to the ESL classroom are discussed. Finally, ideas for future research based on the findings of the present study are offered. The discussion begins with a review of the basic constructs of Output and developing L2 pragmatic competence as they were realized in this study.

** Constructs: Output and Developing L2 Pragmatic Competence 

The notion of Output, as understood in the framework of Swain’s (1995, 2005) Output Hypothesis, involves learners’ actively producing language as part of the process of acquiring a second language (Swain, 2005). Citing Swain’s research, Hall and Verplaetse (2000) emphasized the important role of Output in interaction in the second language classroom setting. In the present study, Output was operationalized as text reconstruction on the part of the + Output group. The focus on accuracy and precision of expression entailed in the text reconstruction process qualify this instructional treatment as “pushed” output (Wesche & Skehan, 2002). The fact that the participants in that group were required to reconstruct the texts of the video vignettes alternately in writing and orally helped to ensure that the mode of output would not be an intervening variable when testing the effectiveness of the output-focused instructional treatment. Also, the bi-modal approach to the treatment task matched the two DCT test instruments, one of which required written responses (the WDCT) and the other, oral responses (the ODCT). The participants in the – Output group were not provided with opportunities for output, but rather were required to answer comprehension or extension questions based on the content of the video vignettes (see Appendix J).

Developing pragmatic competence in a second language includes second language learners’ demonstration of the emergent ability to perceive and produce the target language in pragmatically appropriate ways. It is important, particularly for raters such as those who participated in the present study who must assess the developing L2 pragmatic competence of learners, that grammatical accuracy not be confused with pragmatic appropriateness. However, as Bardovi-Harlig (2002) has noted, it may not be possible to distinguish grammatical accuracy from pragmatic competence completely, particularly when second language learners at different proficiency levels are being tested.
In the present study, developing L2 pragmatic competence was operationalized as learners’ changes in performance (mean score comparisons) on the three measurement instruments: the pragmatic acceptability judgment task, utilized to test learner perceptions of pragmatic appropriateness (PAJT; Garcia, 2004); the written discourse completion task (WDCT; Beebe, Takahashi, & Uliss-Weltz, 1990; Liao & Bresnahan, 1996; Lorenzo-Dus, 2001), designed to assess a learner’s ability to respond appropriately in writing to a given scenario; and the oral discourse completion task (ODCT; Martínez-Flor & Fukuya, 2005), which functioned to test oral pragmalinguistic production.

**Research Question 1**

The first research question focused on the effect of the output-focused video-based instructional treatment on learners’ ability to recognize and produce pragmatic speech acts appropriately. Taken as a whole, the findings presented in Chapter 4 indicate that the + Output instructional treatment had an effect on some aspects of learners’ pragmatic perception and production. For more accurate analysis and interpretation, the question was divided into three sub-questions, one addressing each of the test instruments. A consideration of each of the sub-questions asked helps to clarify the nature of the effect.

**Question 1a**

Research Question 1a asked whether or not the output-focused video-based pragmatic instructional treatment had an effect on adult ESL learners’ perception of pragmatically appropriate speech acts on a pragmatic acceptability judgment task (PAJT). The pretest-posttest performance of the two groups on the PAJT instrument revealed an overall effect for the + Output instructional treatment. Both the + Output and – Output group demonstrated mean gains from the pretest to posttest, with the mean gain for the + Output group significantly higher. Thus, in answer to Research Question 1a, it can be stated with confidence that the output-based instructional treatment had a significant effect (and a large one, at that) with respect to the perception of pragmatically appropriate speech acts. The results for the PAJT are consistent with Izumi’s (2002) finding that + Output instruction benefited learners to a greater extent than analogous – Output (comprehension-focused) instructional treatment for the acquisition of English relativization by adult ESL learners. As in the case of the present study, Izumi obtained
large effect size calculations for the + Output treatment groups as compared to the non-output
groups in that study. It should be noted, however, that the focus of the grammaticality judgment
instrument utilized in Izumi’s study was formal accuracy rather than pragmatic appropriateness.
The analogy is useful nonetheless in terms of the ability to posit an Output effect for recognition
tasks.

It seems likely that the depth of processing claimed by Izumi (2002) to account for the
improvement of the + Output group in that study also accounts for the effect of the + Output
instructional treatment for the PAJT performance in the present study. The output-focused
treatment required the + Output participants to reconstruct the texts of the video vignettes with
attention to both formal accuracy and pragmatic appropriateness. This combination of
grammatical and pragmatic requirements may be thought of as analogous to the integrative
processing to which Izumi, building on Robinson (1995), attributed successful learning. Pushing
learners to produce utterances or written responses that are both grammatically and pragmatically
acceptable may activate the kind of deep processing that Swain (2005) claimed contributes to the
noticing/triggering function of Output as learners actively work to produce target language forms
(e.g., speech acts) and recognize that they are unable to express their meaning precisely.
Importantly, as Swain points out, this function of Output is not to be confused with the simple
opportunity to “practice” a form, but must involve the consolidation of prior knowledge or the
generation of new understanding on the part of learners (Swain & Lapkin, 1995).

**Question 1b**

Research Question 1b explored the effect of the output treatment on the written
production of pragmatic speech acts in response to situational prompts on a written discourse
completion task (WDCT). The participants’ performances on the WDCT from pretest to posttest
indicated no significant effect for instructional treatment for either group, and in fact revealed a
slight decrease in mean rating score among the – Output instructional treatment group
participants. Closer examination of the results for the two speech act item types (directives and
expressives) revealed that there was a significant mean change from pretest to posttest for the +
Output group on the expressive speech act items, but not on the directive items or the WDCT
instrument as a whole.

The lack of a significant effect for Output despite the calculation of a large effect size for
the + Output treatment may be attributable to a possible ceiling effect that was observed in the
raw data for the WDCT directive item response ratings. A ceiling effect occurs when some participants’ scores on the pretest are too high to enable a treatment group to demonstrate significant improvement on the posttest measurement (Vogt, 2005, p. 40). In the data for the WDCT pretest, it was found that 7 members of the + Output group and 4 members of the – Output group scored above 18 on the WDCT pretest for directive items (of a possible 24 points for the directive items: 6 items rated from 1 to 4 by the raters). These high rating scores on the pretest meant that it was difficult for either group, particularly the + Output instructional treatment group, to improve significantly on the posttest.

One possible way to correct for an observed ceiling effect is to conduct a statistical mean comparison using only the scores of those participants who have scored below a given cut-point on a pretest measurement (Marsh, 1988). In the case of the present study, such a special paired-samples $t$-test was not part of the original study design, and so was not conducted as part of the data analysis phase of the study. However, in order to explore the possibility that controlling for the possible ceiling effect could reveal more about the effect of the output treatment, it was decided to conduct a special paired-samples $t$-test that included only data from those participants who had rating scores of 18 or less for the WDCT pretest on the directive item responses. The results of that test help to shed light on the nature of the effect for Output, as they revealed a significant effect for the + Output instructional treatment on the WDCT directives, $t(9) = 4.011$, $p < .01$, two-tailed. The results of this test are offered solely to help account for the large effect size for Output calculated on the WDCT pretest-posttest, and indicate that further investigation of the effect of output-focused treatment on written L2 pragmalinguistic production is warranted.

**Question 1c**

The ability of learners to express speech acts in pragmatically appropriate ways on an oral discourse completion task (ODCT) before and after the instructional treatment was the focus of Research Question 1c. The accuracy with which pseudo-conversational oral tests, such as the ODCT used in the present study or the oral proficiency interview (OPI), capture a learner’s actual conversational competence has been subject to question with regard to validity over the past several years (Johnson, 2001; Lazaraton, 1992). Also, because little research has focused on second language pragmatic comprehension (with the exception of Bouton, 1999, and a few others), there is a lack of standard approaches and instruments for measuring and assessing L2 learners’ ability to perceive accurately pragmatic speech acts (i.e., pragmalinguistic input) in the
second language. The results addressing Research Question 1c seem to indicate that the – Output instructional treatment was actually more effective than the + Output treatment. In fact, both the + Output treatment and the – Output treatment were associated with quite large effect sizes.

The fact that only the – Output group showed significant mean rating score improvement from pretest to posttest on the ODCT may be related to the nature of the task, as there is little in the raw data to suggest any kind of ceiling effect on the pretests. It is not likely that the source of the unexpected result is attributable to the instrument not being based on face-to-face interviews. In fact, researchers have found that tape-mediated or print-mediated instruments designed to measure spoken language can correlate quite well with face-to-face oral interviews (e.g., Stansfield & Kenyon, 1992). Rather, the reality may be that both groups benefited from exposure to the pragmatic video vignettes, but that the directive speech act items did not respond well to the output-focused instructional treatment in terms of spoken production. As will be discussed in the interpretation of the results for Research Question 2c, the expressive items did respond well to the + Output treatment.

Perhaps the participants were not developmentally ready to assimilate the information from the instruction on the directive speech acts, recalling the possibility of a developmental pragmatic sequence suggested by Ohta’s (2001) findings and referred to by Bardovi-Harlig (2002). If such is the case, the exposure to the video vignettes in combination with the comprehension-focused activity performed by the – Output group participants may have been sufficient to help them give spoken responses to the directive speech act item prompts without struggling to incorporate specific forms, as the + Output participants may have felt pressured to do as a result of the output-focused instructional treatment that they had so recently gone through. It is not likely the case that mere exposure to the pragmatic video vignettes alone was sufficient for the acquisition, however. Schmidt (1993) argued that it is not exposure to material that leads to the acquisition of pragmatic and discourse knowledge, but rather the ability to learn and make use of communicative strategies. Thus, it is necessary to look deeper in order to account for the ODCT results. One possibility is the combination of communicative pressure and developmental readiness.

The pressure of spontaneous oral production in combination with the struggle to bring to the surface pragmalinguistic forms that had been practiced could conceivably have worked against the + Output participants, at least on those speech acts for which the learners were not
developmentally “ready” to produce (Salaberry & Lopez-Ortega, 1998). It is reasonable to assume that directive speech acts are acquired later than expressives, as directives involve a greater degree of risk within conversation than expressives do and appear to require more planning on the part of the speaker. Such an explanation would be consistent with Bardovi-Harlig and Hartford’s (1991) discussion of how international students in the U.S. use strategies to deal with moves that place them at risk in interaction with native speakers of English and with Bardovi-Harlig’s (2002) assertion that a learner’s developing sociolinguistic ability (Cohen, 1996) affects her ability to control forms to fit a given pragmatic purpose. The learners from the + Output group, faced with the communicative pressure inherent in oral production may simply have lacked the ability to manipulate the forms that they had learned to fit the pragmatic purpose required by the ODCT directive prompts. The – Output group participants, on the other hand, may have relied more extensively on prior knowledge or L1 pragmatic solutions in response to the directive item prompts, which could conceivably have come across to the raters as less awkward than the efforts of the + Output group participants.

**Research Question 2**

The second research question targeted the two different types of speech act that were tested in the study: directives (advice giving and invitations) and expressives (request refusals and compliment responses). The overall question was whether or not the output-focused instructional treatment had a different effect on directive and expressive speech act item response performance among the participants. The question was motivated by the possibility of a developmental sequence for pragmatic speech acts (Bardovi-Harlig, 2002; Beebe & Waring, 2004; Ohta, 2001). As with Research Question 1, the second research question was divided into three sub-questions so that the effect of the output treatment could be more precisely understood in terms of the participants’ performances for the directive and expressive speech act items on each of the three instruments: the PAJT, WDCT, and ODCT measures.

**Question 2a**

Research Question 2a asked whether or not the + Output instructional treatment had a different effect on the directive and expressive speech act items on the PAJT instrument. Testing of the directive item scores from pretest to posttest revealed a significant effect for Output on the
directive speech act items, and a large effect size as well. The results for the PAJT expressive items, on the other hand, did not reveal any significant effect for Output. However, the pretest data for the expressives exhibited characteristics of a ceiling effect, and the output treatment was associated with a medium effect size, so further exploration of that data was determined to be useful in order to arrive at a better-informed interpretation of the results.

As with the WDCT results (see the discussion of Research Question 1b above), trends in the raw PAJT expressive item response score data and the effect size calculated for the Output treatment revealed an effect for that treatment not identified in the ANOVA, again suggesting a possible ceiling effect in the learners’ pretest performances. In this case, the cut-score for a special t-test designed to look at the possible contributions of the performances of those who did not perform at close to ceiling was set at 4, so that all participants who scored 5 or 6 (of a possible 6) for the expressive items were excluded from the test. Recall that the PAJT responses are either correct or incorrect, rather than representing the mean ratings (from 1 to 4 points) assigned by two outside raters. When the high scorers on the pretest expressives were excluded from the testing, there were 9 participants remaining in the + Output group and 8 in the – Output group. The special t-test was then conducted, and it revealed a significant effect for Output for the expressives, \( t(8) = 3.162, p < .05, \) two-tailed. Again, although this secondary analysis was not part of the original study design, it served the purpose of allowing the researcher to understand the trend evident in the raw data and the effect size calculation. In terms of Research Question 2a, the answer supported by the data analysis is that the PAJT directive and expressive items did respond differently to the + Output instructional treatment, with the directive item responses being more favorably affected than the expressive item responses.

**Question 2b**

Research Question 2b explored the possible differential effect of Output on the directive and expressive speech act items on written production of pragmatic speech acts on the WDCT from pretest to posttest. As noted in the discussion of Research Question 1b above, the participants’ performances on the WDCT for directive speech act items indicated the possibility of a ceiling effect, prompting further exploration of the data in the form of a special paired-samples t-test, which revealed an effect for Output on the directives when the ceiling effect was controlled for. Taken as a whole, however, the data for both the directive and expressive speech act items indicate a greater effect for Output on the expressives than on the directive items. First,
the large effect size calculated for the output effect on expressives was not matched on the directive speech act items. Also, there was a measurable, significant pretest-posttest increase in mean rating score for the + Output instructional treatment group on the expressive items, although the effect for the output treatment was not observed on the directive items without controlling for the possible ceiling effect. It is noteworthy that the directive speech act item responses tended to respond better to the output treatment on the PAJT instrument, whereas on the WDCT instrument, the expressive item responses responded more to the output effect. The reasons for this difference may have to do with the differences between a recognition task (the PAJT) and a production task (the WDCT) as well as with the functional and formal differences between directive and expressive speech acts.

First, recognition tasks such as the PAJT may involve less learner attention to formal features of the target speech acts than production tasks such as the WDCT. It is widely recognized that receptive knowledge (recognition) requires less depth of learning of a given form than productive knowledge does (Nation, 2005). It may be that the + Output group participants were pushed to give more attention to the formal properties of the speech acts they had learned when they were confronted with the WDCT prompts than they were when choosing from among options on the PAJT instruments.

Secondly, the nature of directive speech acts, which are intended to direct another person to take some action in response to the utterance, is different from that of expressive speech acts, which serve to express a speaker’s reaction or attitude in response to another person’s speech or action. Specifically, directives are more likely to be high in communicative pressure because of the fact that they are more likely to be perceived as face-threatening acts than expressive speech acts, which are often employed as face-saving acts, appealing to a listener’s sense of negative face—her need to feel that her time and concerns are valued by the speaker (Yule, 1996).

It may be, then, that the facilitative effect of Output extends more readily to forms with less attached communicative pressure (i.e., the expressives) when learners are required to produce the forms (i.e., on the WDCT or ODCT task), whereas the forms with greater attached communicative pressure (i.e., the directives) are more easily recognized on the less pressure-filled and more meaning-focused recognition task (the PAJT). This would be consistent with Bardovi-Harlig and Salsbury’s (2004) finding that increased opportunities to use the language
led to better performance on disagreement moves. A summary of the possible relationship among the tasks and items appears in Figure 10.

![Figure 10. PAJT and WDCT/ODCT tasks related to speech act types.](image)

**Question 2c**

Research Question 2c targeted the potential different effect for Output on directive and expressive speech act items on the ODCT task from pretest to posttest. If the model in Figure 10 is accurate, it would be expected that the + Output instructional treatment would have a greater effect on expressive speech acts than on directive item responses. In fact, the data did reveal a greater effect for Output on expressives than on directives. For the directive item response rating scores, neither the + Output nor – Output group was associated with a significant effect for treatment, and neither demonstrated a significant change in mean score on the directives from pretest to posttest on the ODCT. As with the WDCT directive results, this finding could be explained by the higher communicative pressure of the directive items in combination with the greater focus on formal features demanded by the ODCT task, which, like the WDCT task, required participants to produce the speech act needed in response to a given prompt.

The effect for Output on the ODCT expressive items was much more evident. First, the raw data for the ODCT expressives revealed a noticeable mean increase from pretest to posttest.
for both the + Output and – Output groups. There was a significant effect for both the Output and non-Output instructional treatments, and very large effect sizes calculated as well. Thus, in answering Research Question 2c, it seems that the lower communicative pressure attaching to the expressive speech act items and the greater focus on formal features associated with the ODCT task combined to facilitate a greater effect for Output on the expressives than the directive speech act items on the ODCT from pretest to posttest.

With respect to the formal and functional differences between directive and expressive speech acts, additional consideration must be given to the fact that directives are generally more syntactically complex than expressives. Bardovi-Harlig (2002) mentioned the form of speech acts as one key way in which learners’ production may differ from that of native speakers. It may be that the “Grammar Halo” posited in the model of pragmatic acquisition (see Figure 1, p. 9) constrains learners’ ability to implement directives effectively on the demanding DCT tasks. Gass (2002) also noted that the “foreigner talk” directed to learners tends to favor more syntactically simple constructions. The nature of the input available to learners could thus also be a contributing factor to the preference for expressives demonstrated by the learners in the present study on the more demanding DCT tasks.

**Research Question 3**

The third research question concerned the internal consistency and reliability of the pragmatic rating rubric used in the present study to assess the responses of the participants on the DCT instruments (the WDCT and ODCT pretest and posttest measures). The results of the statistical calculations for inter-rater reliability and internal consistency indicated that acceptable levels of internal consistency and reliability were observed for the WDCT pretest, the WDCT posttest, and the ODCT pretest. The reliability level calculated for the ODCT posttest, however, was just below what is normally considered to be acceptable in social science research (Vogt, 2005).

Overall, the third research question can be answered affirmatively. That is, the rating rubric as utilized in the present study was sufficiently consistent and reliable as a whole, given the training and practice provided to the non-expert raters prior to the rating process. The only real challenge to the reliability of this finding—the coefficient alpha calculated for the ODCT
posttest—was not sufficiently low to call into question the reliability of the rubric as a whole. Also, the challenge to reliability represented by that lone result was likely the product of one or more practical challenges common to all efforts to assess learners’ L2 speech production. This possibility is explored further in the upcoming discussion of study limitations.

**Summary of Questions and Findings**

In summary, the findings of the study generally support the application of Swain’s (1995, 2005) Output Hypothesis to the acquisition of second language pragmatic ability. Significant effects for Output are clear for pragmatic recognition and may be argued for (based on effect size calculations) for pragmatic production, at least for some speech act types. The possibility of a developmental sequence in second language pragmatic acquisition was supported, although the findings of the present study suggest that the task (or context) may play an important role in how given speech acts are realized. The rating rubric used to assess the pragmalinguistic responses was generally supported, although not completely. An overview of the research questions and the findings that directly address each question facilitates understanding of the effect of the output-focused instructional treatment, the ways in which the + Output treatment had a differential effect on directive and expressive speech act items, and the internal consistency and reliability of the pragmatic rating rubric.

**Limitations**

The limitations of the present study are acknowledged. This section addresses limitations in the areas of rating and the pragmatic rating rubric used in the study, the analysis of the study data, and some practical challenges associated with this classroom-based research effort.

**Rating Issues**

The ability to judge pragmatic responses is a subjective, human endeavor, and so is subject to error. In order to control for this potential source of error, multiple raters were employed for the scoring of responses, and internal consistency and reliability promoted. One possible reason for the lower internal consistency and reliability observed for the ODCT posttest ratings may have been a lack of consistency in the understanding of the rating rubric levels
among the 28 raters. Mackey and Gass (2005) noted that in the field of second language research, there are no accepted standards for rating scales, and the same is true in the subfield of acquisitional pragmatics research. The rationale behind the use of four rating levels was that four levels are enough to distinguish among completely acceptable, generally acceptable, generally unacceptable, and unacceptable pragmalinguistic responses while not overwhelming the raters with an inordinately large number of levels or giving raters a neutral option, or a middle level to which they can default when they have difficulty making a decision about the level to assign to a given response. The rating scale in the present study also followed the suggestion of Mackey and Gass that the scale be based on descriptors. Two different raters, however, may interpret the same verbal descriptor differently. Although such descriptors are preferable to simple numerical rankings, they are not free from problems.

A second possible explanation for the lack of consistency and reliability in the ODCT posttest ratings may be that the raters simply were not familiar enough with the complications that can accompany the evaluation of spoken language for pragmatic acceptability. The need for raters to be trained and prepared appropriately to assess learners’ spoken responses has been well documented. Alderson, Clapham, and Wall (1995) discussed the need for raters to be trained in how to assess language learner test responses accurately, even addressing information on evaluating specific language skills directly to raters. Chalhoub-Deville and Fulcher (2003) reviewed several measurement and rater issues that may contribute to concerns about inter-rater reliability, including a discussion of oral interview assessments as constituting a unique genre of interaction (p. 503).

Focusing on communicative competence, Salaberry (2000) pointed out that the generally accepted notion of communicative language ability often does not recognize the very important role of strategic competence, which may contribute more to success on some interview-style oral language assessments than linguistic ability does. Likewise, Jenkins and Parra (2003) studied the performances of international teaching assistants (ITAs) taking an oral proficiency test and found that those who demonstrated non-verbal behavior and discourse strategies that more closely resembled native speaker behaviors and strategies were rated higher by interviewer/raters when overall language proficiency was neither clearly very high nor clearly very low (pp. 102-103). The training and preparation of raters is essential if they are to develop the awareness of pragmatic and interactional dynamics essential to accurate evaluation.
It may be the case that the raters in the present study were also demonstrating some level of rating fatigue (analogous to the task fatigue discussed by Munro & Mann, 2005). The rating process was carried out over the span of three weeks, following two training sessions in two successive weeks. Because the ODCT posttest responses were evaluated toward the end of the rating process, it is possible that there was a fatigue effect. Alternatively, it could be that the raters continued to move toward their own, personalized ways of thinking about the rating levels over the course of the three weeks so that by the end of the process, they had established a degree of individual understanding of the characteristics of each rating level that adversely affected the ability to rate items consistently with other raters. Some suggestions for ways in the rating issues raised here should help inform future research are offered in the upcoming discussion of recommendations for acquisitional pragmatics and ESL practitioners.

**Data Analysis Issues**

Isolating Output as a causative variable is no easy task. The study followed previous research (e.g., Beebe, Takahashi, & Uliss-Weltz, 1990; Izumi, 2002) wherever possible in order to ensure that the overall design and methodology were consistent with those prior efforts to understand the effects of Output while also addressing questions that had not yet been sufficiently explored. The fact that significant effects for Output were not identified by statistical tests for two tasks (the WDCT overall and ODCT overall) despite observable mean increases from pretest to posttest and large effect sizes calculated for the + Output instructional treatment is indicative of two related limitations of the study in the area of data analysis.

First, the size of the sample for each instructional treatment group \((n = 17\) per group) was not very large. Fraenkel and Wallen (2003) indicated that experimental or quasi-experimental studies in educational settings should include at least 15 participants per treatment group, so the sample size is not inordinately small. However, the number of scores may have been too low for the analysis of variance (ANOVA) to identify the significant effects and interactions suggested by the trends in the raw data and effect size measures. Because the \(F\) value calculated by the ANOVA represents the ratio of the amount of between-group variation to the amount of within-group variation, the larger the sample size per group, the more likely a significant effect or interaction will be identifiable (Mackey & Gass, 2005). This limitation can serve to inform future research; perhaps replication studies could be conducted in multiple locations in order to provide not only larger sample sizes, but also more generalizable results (Creswell, 2005).
The second limitation in the area of data analysis is also related to the sample in each treatment group. As indicated in the discussion of Research Questions 1b and 2a, there was a possible ceiling effect observed in the pretest raw data for some of the participants (7 + Output and 4 – Output participants for the WDCT directives; 8 + Output and 9 – Output participants for the PAJT expressives). In order to explore a possible obscured effect for Output on these tasks and speech acts, two special $t$-tests compared the pretest-posttest means of the learners who did not score above a given cut-point on the pretest. In both cases, an effect for Output was identified by the special $t$-test, but because it is not possible to definitively demonstrate that a ceiling effect existed for the learners who were excluded, these findings were not used to answer the primary research questions. Rather, the possibility of a ceiling effect should help to inform future classroom-based research into speech act perception and production, perhaps indicating that future studies might determine grouping based on participants’ pretest scores, rather than mixing classes of learners at different overall proficiency levels in an effort to control for learner differences, as was done in the present study. This type of design decision would need to be considered carefully, of course, as it would almost certainly involve taking learners out of their usual classroom groups, creating a less authentic environment for the treatment and posttesting.

**Practical Challenges**

The conduct of research in the second language classroom setting is always open to error and practical challenges. The most notable practical challenge that faced the present study concerned the collection of data for the ODCT. As discussed in previous chapters, the oral measure data was collected by having the participants record their spoken responses to the DCT prompts on handheld recording devices or personal computers. The technical problems encountered during the data collection process have been discussed, but a word about how these problems could help to inform future research efforts is in order. Because the problem with data collection that emerged during the present study was a practical problem having to do with the functioning of the recording equipment, the most obvious recommendation would be to establish a system whereby the researcher or an assistant is able to check the recorded responses in real-time (as they are completed by the participants) rather than after the fact. This could be a logistical problem depending on the physical setting of the data collection environment, the number of participants, and the time allotted for the collection of oral responses.
Future Directions for Research

Several of the findings of the present study warrant further investigation in future research efforts. This section provides ideas for the application of the findings to future research through exploring trends in the data, investigating a pragmatic acquisitional sequence, targeting other speech act types, and testing for effects of Output in other areas of L2 acquisitional pragmatics.

Exploring Trends in the Data

Mackey and Gass (2005) noted that second language researchers sometimes use terms such as “approaching significance” or “demonstrating trends” when \( p \)-values for statistical tests fail to meet the usually accepted level of .05 or less for statistical significance, but are less than .10. In fact, accepted statistical philosophy does not encourage exploring or commenting on results that fail to meet the initial alpha value set by the researcher. In other words, if the researcher determines that an alpha of .05 gives sufficient power to reach conclusions in a particular study, it is not recommended to consider any resulting \( p \)-value (in the case of ANOVA) that is above that .05 level (A. Kamata, personal communication, January 24, 2007). In the case of the present study, there were two tasks for which a large effect size was calculated despite the lack of a significant effect or mean change from pretest to posttest identified by the statistical tests for the + Output instructional treatment: the WDCT task and the ODCT task. In both cases, when the individual speech act types were investigated, it was found that there was a significant effect for Output on the expressive speech act items (see Table 48 and Figure 10).

The primary implication for future research here is that the constructs of pragmatic perception and production must be understood as complex, multifaceted phenomena. The design of future studies must target specific aspects of pragmatic perception and production in order for acquisitional pragmatics research to move beyond surface understanding of the effects of various types of instruction on acquisition of L2 pragmatic competence. Also, increasing the number of participants included in future research efforts above that used in the present study, possibly by coordinating testing at multiple sites, could help to identify significant effects and interactions more readily (see Mackey & Gass, 2005). Finally, the use of a measure of effect size, such as Cohen’s \( d \), is encouraged in cases where a smaller sample size or other factors may obscure
effects that may otherwise have been revealed in standard statistical tests such as the ANOVA or t-test (A. Kamata, personal communication, January 24, 2007).

**Pragmatic Acquisitional Sequence**

In order to test more rigorously for the existence of a developmental sequence in L2 pragmatic acquisition, it is important that three things be done: a broader range of speech act types needs to be tested, studies that have as their principal aim the exploration of the developmental sequence need to be conducted, and meta-analyses of speech act-related studies that have implications for the developmental sequence need to be carried out. Each of these important contributors to the establishment and understanding of the developmental sequence in L2 pragmatic acquisition is explored below.

First, a broader range of speech act types than has been tested in past research or in the present study needs to be considered in future acquisitional pragmatic studies. This call for a broadening of the scope of the speech acts covered in future research is in line with Bardovi-Harlig’s (2002) admonition address the central question that must drive SLA pragmatics: “How does L2 pragmatic competence develop?” Exploring the order of acquisition of a greater variety of speech act types helps to address that central question in that such research goes beyond supporting the possibility of a developmental sequence, as the present study has done, to providing more detailed information about the nature of that sequence.

The establishment of a developmental sequence in L2 acquisitional pragmatics should lead to more studies that have as their primary focus the nature of the sequence itself. The investigation of different speech act types and how they are learned and influenced by various instructional treatments must be part of that effort. Most studies that have had implications for a developmental L2 pragmatic sequence have actually focused on other issues, such as Ohta’s (2001) study, which was concerned with how learners of Japanese as a foreign language learned to express alignment with interlocutors, or Bardovi-Harlig and Salsbury’s (2004) study, which focused on learners’ ability to express opposition (disagreement). As mentioned earlier, the syntactic complexity among speech act types is also likely to be an important factor that combines with sociopragmatic awareness to influence learners’ readiness to acquire given speech act types. It is important to note that pragmatic and linguistic factors are not separable, but must be considered together in future research efforts. In order for developmental L2 pragmatics (or acquisitional pragmatics) to progress as a branch of SLA research in its own right, it is critical
that a developmental sequence of speech act acquisition be established and explored as a central object of inquiry.

Research efforts that focus on acquisitional sequences in L2 pragmatics will almost certainly advance the field of acquisitional pragmatics, much as the “morpheme studies” of the 1970s and 1980s (e.g., Dulay & Burt, 1973) spurred the establishment of developmental sequences for L2 morphosyntax and drove the broader field of SLA research forward. Meta-analyses of the accumulated speech act-related acquisitional pragmatic studies can then be carried out to identify significant results across findings in order to improve understanding of the nature of developmental sequences in L2 developmental pragmatics.

Targeting Other Speech Act Types

As indicated above, the testing of different types of speech acts and how they are acquired by L2 learners is essential to a research agenda that can drive acquisitional pragmatics to more systematic understanding of the process of L2 pragmatic development. The present study investigated the effect of an output-focused instructional treatment on directive and expressive speech acts and the results revealed that these two speech act types responded to the instructional treatment and also appeared to respond differently depending on the outcome measurement that was used (i.e., the PAJT and the two DCT tasks). In future research, other speech act types need to be investigated in order to determine if the differential effect of Output (or other instructional treatment) applies more broadly. Two other speech act types that could readily be tested are representatives and commissives (Yule, 1996). As mentioned above, there are social and syntactic factors that may contribute to the facility with which learners can acquire these particular pragmalinguistic forms. In addition, the exploration of a possible developmental difference between direct and indirect speech acts might be a fruitful area of inquiry in future acquisitional pragmatic research, based on the findings of Bardovi-Harlig and Salsbury (2004), Liddicoat and Crozet (2001), Martínez-Flor and Fukuya (2005), and the present study.

Output in Other Areas of Acquisitional Pragmatics

Aside from the need to focus on a variety of speech act types in future research efforts, it is also important to explore the possible effect of Output on other aspects of acquisitional pragmatics. For example, it might be fruitful to investigate how output-focused instruction affects learners’ ability to make off-record statements (i.e., indirect suggestions) as opposed to on-record (direct or imperative) statements. Another potential area of exploration is the efficacy
of Output for helping learners develop the ability to respond to disagreement and potentially tense situations, such as those investigated in Davies and Tyler’s (2005) study of a Korean international teaching assistant in the U.S. who struggled to use appropriate “discourse strategies” (p. 154) when discussing grades and cheating with two American students whom he was teaching. Whatever the targeted outcomes, it is important that output-focused treatments be comparable across studies so that results can be usefully assimilated into the broader effort to explore the efficacy of Output in acquisitional pragmatics.

**Recommendations for Acquisitional Pragmatics and ESL Practitioners**

The following recommendations follow from the study results and discussion. It is important that the results of research, particularly investigations that are classroom-based, be applied to the actual teaching and learning of second or foreign languages. This section of the chapter includes suggestions in the following areas: the use of a didactic model for teaching pragmatics in the L2 classroom; the application of the pragmatic video vignettes to broader contexts; the development of a taxonomy of developing L2 pragmalinguistic forms; the further enhancement and use of the pragmatic rating rubric; and the training and development of raters, particularly non-expert raters.

**A Didactic Model**

Mir’s (2001) three-phase didactic model offers a useful framework within which the findings from the present study could be applied. As was mentioned in Chapter 3, Mir’s model was one of the influences on the design of the instructional treatment for the present study. This proposed model has been adapted here for instruction across four class periods (2 class periods each week, for example) and incorporates elements of the instructional treatment from the study that were indicated to be effective in helping the participants to improve in the perception and production of pragmatic speech acts. The students are taken through the following phases of instruction: a diagnostic phase, a presentation phase, and an analysis phase.

In the diagnostic phase, learners are taught specific forms that can be used to carry out the speech act. Next, in the presentation phase, students are shown a video clip portraying the speech act they have just learned about in a dialog, but without sound. The students next work in pairs to try to create a dialog to match the muted video; their ideas are discussed as a class. Then,
the video clips are re-watched with full sound, so that students can check their ideas against the native speakers’ production of the speech acts. The students are then given transcripts of the video clips to study for approximately one minute each. They then work independently to reconstruct the dialogs orally or in writing. Finally, in the analysis phase, the transcripts are re-read aloud, and the learners are also provided with other handouts that will explicitly teach the speech acts contained on the videos. The students then have a final discussion of the application of the speech acts in different situations and re-watch the video clips one final time.

Based on the findings of the present study, the instructional model outlined above would be expected to improve learners’ recognition of pragmalinguistic forms most immediately. The effects of the instruction on written and oral pragmalinguistic production would possibly depend on the types of pragmatic speech act targeted, as indicated in the findings associated with the second research question in the present study. The effectiveness of such a program of instruction would also be expected to be linked with the overall L2 proficiency of the learners.

**Broader Applications of Pragmatic Video Vignettes**

The pragmatic video vignettes used in the present study also have possible applications in other areas of second language instruction and assessment. First, it is possible to use video-based pragmatic instruction in the U.S. higher education context to help equip international students and teaching assistants (ITAs) for situations that are likely to occur in the university classroom. Such an application would address the call by Davies and Tyler (2005) to focus pragmatic instruction (and research) on discourse strategies, which they define as the purposeful use of language to achieve communicative goals (p. 154).

Extensions of the use of the pragmatic video vignettes are suggested in the work of Kinginger (1998, 2000) on videoconferencing in the French as a foreign language classroom and in Nedashkivska’s use of video clips to teach pragmatic elements to learners of Ukrainian. The use of video-based instruction to teach and reinforce the understanding of speech acts in preparation for video interaction with native speakers of a target language has obvious potential for use in the second language classroom. Learners who are better prepared for interaction with conversation partners will likely be in a better position to benefit from the interaction in terms of their continued L2 pragmatic development.

A final possible use of the pragmatic video vignettes discussed here is for assessment purposes. As indicated in the didactic model outlined above, based on Mir’s (2001) work, it is
possible to use the video vignettes to not only teach pragmatic speech acts, but to assess second language learners’ ability to supply appropriate speech acts in given situations. Learners in the ESL or FL classroom may be instructed to view the muted video vignettes, as indicated in the didactic model, but rather than working with a partner to construct an acceptable speech act, each learner is charged with supplying a speech act on her own, either orally or in writing. As Leung (2005) suggests, such ongoing assessment can be an effective tool for promoting learning provided it is integrated with the everyday teaching and learning process at work in the classroom.

A Taxonomy of Developing Pragmalinguistic Forms

Kasper and Rose (2002) have posited, based on a review of several longitudinal studies of developmental pragmatic ability, that early-stage learners acquiring a second language tend to rely on unanalyzed formulae, gradually learning to use of productive pragmatic elements in their speech (pp. 125-132). This observation suggests that raters or teachers who are attempting to assess second language proficiency need to take the learner’s developmental stage into account. It may not be possible for beginning or even intermediate level learners to adapt spontaneously a known pragmatic formula (e.g., “no, thank you”) to an unfamiliar situation. In response to these realities, there could be great benefit in the creation of a taxonomy of pragmatic constructions (pragmalinguistic forms), possibly coordinated with overall L2 developmental levels, such that raters could consult the list when making decisions about an ELL’s performance on an assessment instrument. The creation of a taxonomy such as that described here would involve continued research into how pragmatic speech acts are acquired as well as thorough meta-analysis of studies, such as that conducted by Norris and Ortega (2000) for general SLA.

The Pragmatic Rating Rubric

The rating rubric utilized by the raters to assess the oral and written pragmatic responses of the participants on the DCT instruments in the present study could be applied in a variety of measurement and evaluation settings where pragmatic competence is an area of focus. Based on the results of the present study (Research Question 3), the major area of concern is the usefulness of the rubric with respect to assessing oral L2 pragmatic speech act production. The general suitability of rating scales for evaluating pragmalinguistic production has been touted by Kasper and Roever (2005), among others. The key to successful implementation of a rating rubric, however, is not solely in the appropriate application of the scale, but in the careful design and
pilot testing of the rubric as well. In the case of the proposed rubric used for the present study, the process of testing the rubric for reliability is already well underway. What is needed is careful scrutiny of each of the levels of the rubric in order to determine if there is any inconsistency or ambiguity that may have contributed to the difficulty that the raters had in reaching consistent scoring decisions for the ODCT posttest responses. It has already been noted that oral responses are more difficult to assess than written responses, but identifying the source or sources of that difficulty when pragmatic speech act production is the objects of evaluation is critical to the successful development and implementation of the pragmatic rating rubric.

**Pragmatic Rater Training**

The results of the third research question revealed that the rating rubric as utilized in the present study was generally consistent and reliable. Certainly one reason for that generally positive result was the training provided to the non-expert raters before beginning the rating process. A possible useful application of the findings of the study is the expansion of rater training for the purpose of assessing L2 oral language to include greater focus on pragmatic and sociolinguistic elements.

Even when non-expert raters are given practice in listening to and rating language learners’ spoken responses, they may not be prepared to deal with aspects of spoken language beyond grammatical accuracy and general intelligibility. Young’s (2002) analysis of several studies comparing oral language tests found that raters tended to focus primarily on grammatical competence and fluency when assessing lower-level students in interview-type assessments, but for higher-level students, the use of “stylistic devices” was viewed as an important feature of learner responses (p. 253). Adding further support to the call for raters to be trained to recognize the importance of pragmatic factors and communicative contexts, Bardovi-Harlig and Hartford (1991) suggested that the maxim of congruence is at work in talk between native speakers and non-native speakers. This maxim holds that one’s contribution to a conversation should be congruent with one’s status relative to the interlocutor in a given situation, or context. Bardovi-Harlig and Hartford added a corollary to the maxim: if conversational congruence is not possible, a speaker should use status-preserving strategies to mitigate the negative effects of non-congruence (p. 42).

Raters of learner language samples, such as the responses elicited in the present study, need to be aware of the dimensions of pragmatic acceptability that are necessary to a more or
less appropriate learner response on an instrument such as the discourse completion task. The training provided to the raters in the present study can readily be expanded to include more extensive explanations of the various strategies and maxims at work in the pragmalinguistic responses of learners. Such preparation will be helpful if care is taken to expose the raters to sufficient amounts of authentic data exemplifying the strategies and maxims and representing the range of rating levels of the scale being used. A well-tested pragmatic production rating scale would be an important tool for use in both acquisitional pragmatic research and in the second language classroom for assessment and instructional purposes.

Concluding Thoughts

The present study has provided evidence for the effect of an output-focused instructional treatment in the context of L2 pragmatic development. As such, it offers support for Swain’s (1995, 2005) Output Hypothesis in an area of second language development outside of morphosyntactic learning. Importantly, the study also offered evidence for a differential effect for Output on two different types of speech act (directives and expressives). This finding suggests the need to explore more deeply the ways in which different speech act types may respond to output-focused pragmatic instruction. The findings of the study further suggest that the effects of Output may be linked to the type of task that is used to measure pragmatic perception and production—a finding that could lead to groundbreaking research into how learners balance the need to focus on formal features of pragmatic speech acts with the need to work under the communicative pressure that certain speech acts entail. The generally high level of internal consistency and reliability found for the rating rubric means that this instrument has the potential to be adapted for future research and even classroom assessment purposes. Finally, the innovative use of video-based pragmatic instruction in the present study should be pursued not only in research settings, but in the second language classroom in order to teach more successfully the complexities of pragmatics. It is hoped that the present study will contribute to the growing body of research at the intersection of acquisitional pragmatics and second language instruction from which it has grown, and that future research will be able to draw from and build upon the findings presented here.
Pragmatic Acceptability Judgment Task – Pretest

For each dialogue, choose the option (a – d) that the speaker should use to respond most appropriately to the situation.

Situation 1:
A student is attempting to pay his housing bill, but there is a problem when he realizes that he doesn’t have his student ID with him.

Woman: Hi, can I help you?
Student:  Hi.  I need to pay my summer housing bill—for my dorm room.  There should be something due today?
Woman: Sure.  Could I see your student ID, please?

Student:    a) Actually, I can’t show my ID to you.
           b) Sorry, but I don’t have my ID with me right now.
           c) Sorry, I can’t show my ID card at this time.
           d) Sorry, I can’t do that, OK?

Situation 2:
A student is talking to his professor about an upcoming test in the class.

Student: Dr. Jones, I feel I’ve prepared pretty well for Friday’s exam, but I still have a few questions.
Dr. Jones: Sure, what can I help you with?
Student: Well, I’m wondering which part of chapter four I should focus on most.

Dr. Jones: a) Actually, you must study all of chapter four, OK?
           b) You should study the entire chapter if you know what’s good for you.
           c) Why don’t you study all of chapter four?
           d) You really should be prepared for questions from all of chapter four.

Situation 3:
Two female friends meet each other before their university class begins.

Susan: Hey, Sarah. How was your weekend?
Sarah: Pretty good. We spent most of Saturday working on the house.
Susan: Say, is that a new sweater?
Sarah: Yeah. My only blue sweater had a hole in it, so my sister bought this one for me.
Susan: It looks really great!

Sarah: a) I know, it’s cool.
     b) Thank you, I think it looks great on me, too.
     c) Thanks! I was really happy when she gave it to me.
d) No, it’s really not so nice.

**Situation 4:**
A mother is talking to her son on the phone about a birthday party that is coming up for a relative.

**Mother:** Hi son, I’m calling to let you know about a birthday party this weekend.
**Son:** Hi Mom. Who is the party for?
**Mother:** Do you remember your uncle Robert?
**Son:** Yes. I remember Uncle Rob.

**Mother:**
- a) I want you to come to the party, so please be there.
- b) You are welcome to come to the party if you’re not too busy.
- c) I would like to invite you to attend the wonderful party this weekend.
- d) You are hereby invited to come to the party for your uncle.

**Situation 5:**
A professor is talking to one of her students about the possibility of the student participating in an upcoming conference at the university.

**Professor Sanders:** You’ve been doing very well in the class, Jeff, and I think you would really benefit by participating in the upcoming conference.
**Jeff:** Thanks, Dr. Sanders. What would I need to do?

**Professor Sanders:** Why don’t you come to the meeting room early—at around four o’clock—and help the other students set up the stage and the welcome table.

**Jeff:**
- a) I’m sorry, but I’m not coming that early.
- b) I’d like to help, but at such an early time, I cannot participate.
- c) I would help if I could, but I am not able to participate at that early hour.
- d) I’d like to help, but I’m afraid I can’t make it that early.

**Situation 6:**
A student is talking to her teacher, who is a graduate teaching assistant, about the grade she received on a recent quiz.

**Student:** Hi, Angela, I want to ask you about my grade on the quiz.
**Teacher:** Sure, Sarah. What’s your question?
**Student:** Well, you marked two of the questions wrong when I’m pretty sure my answers are possible. I mean, they could be correct, depending on how you read the questions.

**Teacher:**
- a) Oh, but for these kinds of questions, you really should respond based on the context that I give you on the paper.
- b) Yes, but you cannot answer these kinds of questions without the right context that I give you on the paper.
- c) Oh, but you should answer these kinds of questions in a different way, don’t you see?
d) Well, you should know that your answers to these questions were not acceptable given the context.

**Situation 7:**
A young man has just given a very good speech in front of a large audience. The audience applauds, and as he takes his seat, the man seated next to him (a friend, but not a close friend) compliments him on his speech.

**Friend:** Were you nervous?
**Speaker:** A little, but I felt pretty good.
**Friend:** Well, you did a great job. I thought the speech was powerful.

**Speaker:**
a) Thanks, I really appreciate that.
b) Do you really think so?
c) I wish it had been good, but I don’t think so.
d) Thanks, I thought it was great, too.

**Situation 8:**
Two students are talking to each other after a lecture class.

**John:** Hey Susan, did you understand what Dr. Smith was talking about at the end of the lecture?
**Susan:** Yeah, he just wanted to make it clear that any of the material covered this week could be on the test next Friday.
**John:** I don’t know; all of this new stuff is just so confusing to me.

**Susan:**
a) If you want, you can join our study group, don’t you think?
b) Well, would you like to join our study group?
c) Well, I would like to invite you to join our study group.
d) Come on and join our study group, please.

**Situation 9:**
Two male students are discussing their plans for the upcoming weekend.

**Ray:** Do you know what you’re doing this weekend?
**Mike:** I think we’re going to the game, then maybe to see a movie on Sunday.
**Ray:** Do you know if there are still tickets available for the game?
**Mike:** I’m pretty sure there are. The team we’re playing isn’t that great.
**Ray:** Do you think you could pick up an extra ticket for my cousin? He’s coming into town on Friday.

**Mike:**
a) Unfortunately, that’s not going to be possible.
b) I really wish I could, but we’ve already picked up the tickets.
c) Well, I would like to help you, but I’m sorry.
d) Oh, sorry, I should do that, but could you understand if I can’t help?
Situation 10:
A father and his son are discussing the son’s plans for the summer.

Father: So, have you decided what you’re going to be doing this summer?
Son: Well, when my classes have finished, I’m thinking about doing some traveling—maybe go to see the Grand Canyon.
Father: Don’t forget that you promised to help your uncle John paint his house this summer, too.
Son: Do you think I should go help Uncle John first and then take my trip? I’m not sure if I’ll have time later this summer. What’s your advice?

Father: a) It should be better for you if you go ahead and help Uncle John first.
b) I hope you will help your uncle before you go on the trip.
c) I think it would be better for you to help Uncle John first.
d) I really think you must help Uncle John, if you want to know my opinion.

Situation 11:
A well-known football player is talking to a young man who is a fan, and who really admires the player’s skill.

Fan: It’s great to meet you. I’m a big fan of the team.
Player: It’s my pleasure. I’m always happy to meet a fan.
Fan: Well, I really think you’re the best player in the league this year. The game last week, you made some amazing plays.

Player: a) Thank you. I believe you.
b) Really? I think last week’s game was a little below average.
c) Thanks. It takes a lot of teamwork to make good plays.
d) That’s a nice thing to say. You’re too kind.

Situation 12:
A woman is staying at a nice hotel. As she walks through the lobby, she meets the director of activities for the hotel. He invites her to join a hotel activity.

Director: Good morning, ma’am.
Woman: Good morning.

Director: a) The hotel is taking a group on a trip to the museum today. You’d better come with us.
b) Would you join a trip that we’re taking to the museum today?
c) You should know that a group from the hotel is going on a trip to the museum today. Could you join us?
d) I’d like to let you know that the hotel is taking a group on a trip to the museum today. You’re welcome to join us.
APPENDIX B

Pragmatic Acceptability Judgment Task – Posttest

For each dialogue, choose the option (a – d) that the speaker should use to respond most appropriately to the situation.

**Situation 1:**
A professor is talking to a student in his class about a paper that the student has written. The professor wants the student to share his paper with the class, but the student does not feel prepared.

**Student:** Hi, Dr. Smith. You wanted to see me about my paper?
**Professor:** Yes. Your paper is well written. I was wondering if you would be willing to share what you have written with the class.

**Student:**
- a) Oh, I’d like to, but I don’t want to do that.
- b) Oh, I’d like to, but I won’t because I don’t feel prepared.
- c) Oh, I’d like to, but I really don’t feel prepared.
- d) Oh, I’d like to, but since I don’t feel prepared, I can’t.

**Situation 2:**
Two sisters are talking to each other about what to wear to class that day.

**Sarah:** Hey Jennifer, are you going to wear your blue sweater to class today?
**Jennifer:** No, I think I’ll wear my yellow blouse.
**Sarah:** Do you think I should wear my new outfit?

**Jennifer:**
- a) Why don’t you wear your green dress instead?
- b) You need to wear your green dress instead.
- c) You must wear your green dress instead.
- d) Why shouldn’t you wear your green dress instead?

**Situation 3:**
Two male classmates are talking to each other after class.

**John:** Are you going to play basketball with the other guys today?
**Bill:** I’m thinking about it.
**John:** Well, you played a great game last week.

**Bill:**
- a) I don’t agree, but thanks.
- b) Thanks, man. I appreciate that.
- c) Oh, that’s OK.
- d) You’re very nice to say that.
Situation 4:
A young woman is talking to her good friend on the phone. She wants to invite her friend to a very nice party that she is hosting.

Mary: Hi Karen, how are you?
Karen: Great. What’s new?

Mary: a) Actually, I was hoping you would attend a party I’m hosting, O.K.?
b) Well, I want to tell you to attend my party.
c) Actually, I want you to come to my party, please.
d) Well, I called you to see if you wanted to come to a party I’m hosting.

Situation 5:
A supervisor at work is talking to a younger man who works for him. The supervisor is asking the worker to help with a project in another building. However, the worker’s keys do not open the door to the other building.

Supervisor: Robert, would you mind helping with our new project?
Worker: Yes, sir. Where are they working?
Supervisor: The team is meeting in Building B.

Worker: a) Oh, I’m sorry, but I won’t go to Building B.
b) Oh, I’m sorry, but my keys don’t work for Building B.
c) I’m sorry, but could you allow me not to go there?
d) Unfortunately, I cannot enter Building B at this time.

Situation 6:
A college student is talking to his teacher, who is a graduate teaching assistant, about the requirements for an assignment in the class.

Student: Hi, David. I want to ask you about the requirements for the assignment.
Teacher: Sure. What can I help you with?
Student: I wanted to know if I should prepare my paper according to the guidelines you gave us last week in class, or using the university guidelines.

Teacher: a) Well, for this assignment, just use the guidelines I gave you in class.
b) Well, for this assignment, you can use the guidelines I gave you, is that clear?
c) Well, why can’t you just use the guidelines I gave you in class?
d) Well, you could try using the guidelines I gave you for this assignment.

Situation 7:
A woman is presenting some information to a group at a business meeting. When she concludes her presentation, the president of the company compliments her on her research.

Woman: Thank you for the opportunity to present my research.
President: That was an excellent presentation, Susan. Thank you.

Woman:  a) Oh, it wasn't really very good.  
b) Do you really think it was excellent?  
c) Oh, thank you, sir. I appreciate that.  
d) Well, I am glad you finally noticed my hard work.

Situation 8:
Two university students are talking to each other just before a lecture class begins.

Jennifer: Rob, what are going to be doing after class today?  
Robert: I don’t know. I’ll probably just hang out at my place for a while.

Jennifer:  a) A few of us are going out to eat, and you may come along if you want.  
b) A few of us are going out to eat, if you care. Want to join us?  
c) Do you want to go out with a few of us to eat or not?  
d) If you want, you could join a few of us who are going out to eat.

Situation 9:
A father is talking to his son, who is a university student, about the possibility of visiting him.

Father: Hi son. Your mother and I are thinking about driving up to visit you this weekend.  
Son: Oh, that’s cool, dad.  
Father: Do you think you could meet us at the student union for lunch tomorrow?

Son:  a) Oh, Dad, I’d like to, but that’s not possible.  
b) Oh, Dad, I’d like to, but I have class until five o’clock.  
c) Oh, Dad, I’m sorry, but I’m not going to meet you then.  
d) Oh, Dad, sorry, but I’ll be busy then, do you see?

Situation 10:
A mother and daughter are discussing the daughter’s plans for finding a new job.

Mother: Hi, honey. Have you made any decisions about the new job?  
Daughter: Not really. There are three companies that I’m interested in, but I’m not sure how to choose.  
Mother: Do you have all of the information about all three companies?  
Daughter: Yes. I have paperwork from all three.

Mother:  a) Well, I think you should compare the companies and just go with the one that seems best to you.  
b) Please just compare the companies and go with the one that seems best to you.  
c) Would you just compare the companies and go with the one that seems best to you?  
d) How about comparing the companies, then you could go with the one that seems best to you, alright?
**Situation 11:**
A famous musician is visiting the music school at a university, and a student in the class approaches the musician to talk to him.

**Student:** It’s an honor to meet you!
**Musician:** Oh, thanks, but I’m just a musician, just like you.
**Student:** Well, your last album was probably the best album I’ve listened to in years.

**Musician:**
- a) Hey, I think you’re right!
- b) No, it wasn't really that good.
- c) I thank you for such a wonderful, generous compliment.
- d) Thanks, man. It was a labor of love.

**Situation 12:**
A man and woman are eating dinner at a very nice restaurant. While they are eating, the waiter comes up and invites the couple to enjoy a bottle of wine or some dessert.

**Waiter:** How is everything?
**Couple:** Very good, thanks.

**Waiter:**
- a) Please take some wine or dessert.
- b) Here, why don’t you have some wine or some dessert?
- c) Would you care for some wine or maybe some dessert?
- d) Why not try some wine or some dessert, if you would.
APPENDIX C

Written Discourse Completion Task – Pretest

(Beebe, Takahashi, & Uliss-Weltz, 1990; Liao & Bresnahan, 1996; Lorenzo-Dus, 2001)

Please respond to each situation as you would in actual conversation

1. You are the owner of a bookstore. One of your best workers asks to speak to you in private.

Worker: As you know, I’ve been here a little over a year now, and I think you’ve been pleased with my work. I really enjoy working here, but honestly, I really need an increase in pay.

You: ___________________________________________________________________
____________________________________________________________________
____________________________________________________________________

Worker: Then I guess I’ll have to look for another job.

2. You are talking to your younger brother about a decision that he has to make.

Brother: You know I’ve always respected your advice. Yesterday, my boss asked me if I wanted to take a new position in the company. I would make more money, but I would also have to move far away. I’m just not sure what to do. What’s your advice?

You: ___________________________________________________________________
____________________________________________________________________
____________________________________________________________________

3. You just got your hair cut in a different style. When you see your friend later that day, he says: ‘Hey, your new haircut looks great!’

You: ___________________________________________________________________
____________________________________________________________________
____________________________________________________________________

4. You want to invite your close friend to attend a party at your house this weekend. What would you say to invite your friend to your party?

You: ___________________________________________________________________
____________________________________________________________________
____________________________________________________________________

5. You are a junior in college. You attend classes regularly and take good notes. Your classmate often misses class and asks you for the lecture notes.

You: ___________________________________________________________________
____________________________________________________________________
____________________________________________________________________
Classmate: Oh no! We have an exam tomorrow but I don’t have the notes from last week. I am sorry to ask you this, but could you please lend me your notes again?

You:  
_____________________________________________________________________
_____________________________________________________________________

Classmate: Oh, well I guess I’ll have to ask somebody else.

6. While you are walking on campus near a parking lot, an older woman, about the age of your grandmother, stops her car and asks you if you can help her.

Woman: Excuse me, I’m visiting campus and was told to park in this parking lot. But there seem to be no available spaces here. Do you have any suggestions about other places to park?

You:  
_____________________________________________________________________
_____________________________________________________________________

7. You are the manager of a company. One of your employees mentions that he is thinking about buying a new car. You have just bought a new car yourself, so you offer to let him take a look at your car. When he sees the car, he says, ‘It looks great! I love this model, and you’ve got great taste in choosing the color!’

You:  
_____________________________________________________________________
_____________________________________________________________________

8. You are a member of the English club on campus. You want to invite your professor, Dr. Smith, to attend a meeting of the club next week. How would you invite Dr. Smith?

You:  
_____________________________________________________________________
_____________________________________________________________________

9. You are the president of a printing company. A salesman from a printing machine company asks you to meet him at an expensive restaurant, but you do not want to.

Salesman: We have met several times to discuss your purchase of my company’s products. I was wondering if you would join me as my guest at the restaurant in order to discuss a contract?

You:  
_____________________________________________________________________
_____________________________________________________________________

Salesman: Maybe another time.
10. You are studying for a big test with a classmate. Your classmate tells you that she has not been able to sleep because of stress. She asks you what she should do.

Classmate: I feel so stressed because of this exam, I can’t even sleep at night. What do you think I should do?

You: _______________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

11. You are a teacher at a vocational school and one evening, you invite your adult students over to your house for dinner. After eating, one of your students says, ‘I didn’t know you were such a good cook. The food was wonderful!’

You: _______________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

12. You have helped to organize a big festival to raise money for a charity. You are asked by the other organizers to make an announcement on the radio inviting the public to come to the event. How would you invite the public to attend the festival?

You: _______________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
APPENDIX D

Written Discourse Completion Task — Posttest

Please respond to each of the following situations as you would in actual conversation

1. You are working at a large accounting firm. One day the boss calls you into his office.

**Boss**: Next Sunday my wife and I are having a little party. I know it’s short notice, but I am hoping all my top executives will be there with their wives. What do you say?

**You**: __________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

**Boss**: That’s too bad. I was hoping everyone would be there.

2. You are talking to your younger sister about an important decision that she has to make.

**Sister**: I’m glad I was able to call you. I have to decide whether to accept a job offer from a really big company or stay in school to get my master’s degree. Do you think I should take the job or stay in school?

**You**: __________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

3. You have just finished a tennis match with a colleague. You have been working hard to improve your tennis game for the last few months, and you win the match. Your colleague says to you, ‘I’d say all your effort has been worthwhile. You played an outstanding game!’

**You**: __________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

4. You want to invite your grandmother to attend your graduation ceremony this weekend. You call her on the phone to invite her.

**You**: __________________________________________________________
_____________________________________________________________________
_____________________________________________________________________

5. You’re at a friend’s house watching TV. Your friend offers you a snack.

**You**: Thanks, but no thanks. I feel like I’ve been eating too much, and I feel just terrible. Some of my clothes don’t even fit me.
Friend: Hey, why don’t you try this new diet I’ve been telling you about?

You: ________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

Friend: You should try it anyway.

6. You are a university student who is working as a waiter in a restaurant to pay for your tuition. While you are working one day, a customer at the restaurant asks for your advice about something to eat.

Customer: Excuse me, what would be your recommendation for dessert?

You: ________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

7. A classmate asks to look at an essay that you wrote for another course at the university and for which you got a very good grade. When he looks over the paper, he says: ‘You really wrote this in a clear and organized way. I wish I could write something this good!’

You: ________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

8. You are a teacher at a local high school, and you are talking to the parents of some of your students. You want to invite the parents to attend a parent-teacher conference next week.

You: ________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

9. You’re at your desk trying to find a report that your boss just asked for. While you’re searching through the mess on your desk, your boss walks over and makes a suggestion, but you don’t want to do what he requests.

Boss: You know, maybe you should try and organize yourself better. I always write myself little notes to remind me of things. Why don’t you give it a try!

You: ________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
**Boss:** Well, it’s just an idea.

10. You are helping one of your colleagues complete an important project. Your co-worker tells you that he does not believe he can complete the project in time. He asks for your advice.

**Colleague:** I really don’t think I can finish this project by the deadline. What should I do?

**You:**

______________________________________________________________________
______________________________________________________________________
______________________________________________________________________

11. Your boss at work has organized a retirement party for one of your co-workers. You dress up in your best suit for the occasion, and when you arrive, one of your other co-workers says, ‘That suit looks really nice!’

**You:**

______________________________________________________________________
______________________________________________________________________
______________________________________________________________________

12. You are working at the local library. A famous author will be coming to the library next week to give a lecture, and you are asked by your supervisor to make an announcement on a local television station to invite people in the community to attend the author’s lecture. How would you invite the public to attend the lecture?

**You:**

______________________________________________________________________
______________________________________________________________________
______________________________________________________________________
APPENDIX E

Oral Discourse Completion Task – Pretest

Please read each situation carefully, then respond to the prompt or question (in bold) by speaking your answer as you would in actual conversation.

1. You’ve been working in an advertising agency for some time now. The boss offers you a raise and promotion, but it involves moving to another, smaller, city. You don’t want to move.

Today, the boss calls you into his office and says:
“I’d like you to consider an executive position in our new office. It’s only about 3 hours from here by plane. And, a nice raise comes with the position.”

How would you refuse his request?

2. Your classmate is thinking about buying a new computer. She isn’t sure about what type of computer to buy.

She says: “I’m not sure whether to buy a laptop or a desktop computer. Also, I am wondering if it’s better to buy a PC or a Mac. What do you think?”

How would you advise your classmate?

3. Your mother comes to visit you in your new apartment.

She walks in and looks around and says, “Wow! You’ve really made the place look so nice. I love the colors and furniture.”

How would you respond to your mother’s compliment?

4. You and your roommates are going to attend a concert on the weekend. You want to invite your new neighbor to go with you to the concert. You only met your new neighbor last week, so you do not know them very well.

How would you invite your new neighbor to attend the concert?

5. You are working at the library on campus. A student wants you to check out some books for her, but you see that her library card is not valid, so you cannot do that.

She asks you, “Excuse me, could you please check out these books for me?”

How would you refuse her request?
6. Your close friend is going to be looking for a new job this summer. He isn’t sure whether he should send out his resume and recommendation letters to the companies he is most interested in, or if he should visit those companies personally.

**How would advise your friend?**

7. You submitted a paper to your professor yesterday, but you are not sure if the paper is good enough to earn a good grade. You visit your professor in his office and ask him what he thinks of the paper.

Your professor says, “I think your paper is very good actually. You answered the questions thoughtfully, and your writing is very clear. I’ll give you your final grade tomorrow, but I would say that you did quite well.”

**How would you respond to your professor’s compliment?**

8. Your father calls you to tell you that your family is going to be having a family reunion next month. Many of your relatives will be coming from far away to this big get-together. He asks you to call your cousin, Joe, to invite him to the family reunion. Joe lives in another state, and you have not seen him in a long time.

**How would you invite your cousin to the family reunion?**

9. You receive a phone call from someone you don’t know. The man on the phone asks you to give him some information, but you are not sure that you can trust him.

Caller: “Hi, I’m collecting information for a survey. Would you be willing to answer some questions about your shopping preferences?”

**How would you refuse the caller’s request?**

10. You are teaching a class at the university, and one of your students approaches you after your lecture to ask for some help.

Student: “I am trying very hard to understand the material in this class, but I’m still confused. What do you think I should do to try to understand the material better?”

**How would you advise your student?**

11. You are playing basketball with some friends, and you take a very difficult shot and you make the shot to win the game. Your friends congratulate you.

Friends: “That was a great shot! You really did great!”

**How would you respond to your friends’ compliments?**
12. You have met someone in one of your classes that you like very much and would like to get to know better. You want to invite this person to have dinner with you so that you can talk and get to know each other better.

How would you invite this person to have dinner with you?
APPENDIX F

Oral Discourse Completion Task – Posttest

Please read each situation carefully, then respond to the prompt or question (in bold) by speaking your answer as you would in actual conversation.

1. You are a teacher and a student asks you if he can submit his homework assignment late. You have told the class that no assignments will be accepted late.
   
   **How would you refuse his request?**

2. A close friend comes to you and tells you that he is thinking of breaking up with his girlfriend. You think that he should not make that decision so quickly.
   
   **How would you give him this advice?**

3. Your aunt and uncle are visiting you from out of town. When they arrive at your home, they look around the house and say, ‘Hey, the house looks terrific! You’ve done an excellent job fixing it up.’
   
   **How would you respond to your aunt and uncle’s compliment?**

4. You want to invite a classmate to a party that you are hosting tomorrow night. She is a new student, so you don’t know her well.
   
   **How would you invite her to the party?**

5. You are working as a waiter in a restaurant. A customer asks you if you can bring him a bottle of wine, but he appears very young and does not have any identification to prove he is old enough to buy the wine.
   
   **How would you refuse his request?**

6. Your classmate is doing poorly in the course that you are both taking. She asks you if you have any suggestions for how to prepare for an upcoming exam.
   
   Classmate: “I’m really worried about my grade in this class. I need to do well on this exam, but I’m not sure how to prepare for the test.”
   
   **How would you advise your classmate to prepare for the exam?**

7. Your neighbor, whom you met only two weeks ago, notices your new jacket and says, “Hey, that’s a nice jacket!”
How would you respond to your neighbor’s compliment?

8. You receive a phone message from your mother asking you to call your sister. Your mother’s message states:

“Hi, this is your mother. I want you to give your sister a call to invite her to your father’s surprise birthday party. I tried to call her myself, but I couldn’t reach her. Just be sure you invite her to Dad’s birthday party.”

How would you invite your sister to the birthday party?

9. While you are walking on campus, a man and woman you don’t know approach you with some papers. They smile and ask you: “Hi, would you mind taking a survey for us? It will only take a few minutes.” You don’t want to do this.

How would you refuse the request of the man and woman?

10. You are a professor at the university. After class one day, one of your students approaches you, looking upset.

Student: “Professor, I know that my performance in your class has not been good. I am thinking about dropping the class. What do you think I should do?”

How would you advise your student?

11. You have just given a presentation in front of your class. After you finish your presentation, you sit down, and your classmate compliments you on your speech.

Classmate: “Hey, nice job on your presentation!”

How would you respond to your classmate’s compliment?

12. You are going to be playing in a soccer match with several of your friends. A lot of people are going to come to watch the match, and you want to invite your roommate to come along.

How would you invite your roommate to come to the soccer match?
APPENDIX G

DISCOURSE COMPLETION TASK RATER TRAINING

Your task:

You will read the responses of an English language learner (ELL) who is responding to twelve different situational prompts. The responses will be one of four types:

1. Request refusals
2. Advice giving
3. Compliment responses
4. Invitations

You must determine how appropriate the ELL’s response is. An appropriate response is not just grammatical and understandable, but is natural given the situation.

To receive a 4, the response should be completely acceptable and natural given the situation. There may be some grammatical mistakes, but these are small and do not interfere with the message at all. A 4 response is the type of response that a native speaker of English might give (or very close to it). Here is a summary of the possible scores:

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Indicators</th>
</tr>
</thead>
</table>
| 4     | response is completely acceptable pragmatically given the context, not noticeably affected by any errors | - approaches native-like usage  
- minor grammatical errors do not interfere with pragmatic effectiveness  
- totally appropriate to the situation |
| 3     | response is generally appropriate given the context, but contains one or more pragmalinguistic flaws that affect the intended meaning | - near native-like usage  
- minor grammatical errors may distract from pragmatic effectiveness  
- may be too brief or too long  
- somewhat appropriate to the situation |
| 2     | response is generally unacceptable pragmatically in this context, though perhaps not in all contexts | - generally non-native-like usage  
- noticeable errors distract from pragmatic effectiveness at times  
- too brief or too long  
- generally inappropriate to the situation |
| 1     | response is unacceptable pragmatically given the context | - clearly non-native-like usage  
- numerous errors distract from pragmatic effectiveness throughout  
- clearly inappropriate to the situation |
Remember that a response that seems too abrupt, rude, or otherwise inappropriate to the situation must be knocked down in points, even if it is grammatically correct. Here are some example responses at each of the four point levels:

<table>
<thead>
<tr>
<th>Score</th>
<th>Example Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Oh man! I am really sorry because I wasn't in that class, too. Maybe you need to ask somebody else. [request refusal; borrowing class notes]</td>
</tr>
<tr>
<td>3</td>
<td>Thank you, man! To do these, it took 8 hours. [compliment response; new hairstyle]</td>
</tr>
<tr>
<td>2</td>
<td>Thank you. I payed [sic] a lot of money for my hairstyle. What hairstyle do you like? [compliment response; new hairstyle]</td>
</tr>
<tr>
<td>1</td>
<td>Thank you to invite me. The food was really good. Could you invite me next time? [compliment response; good cook]</td>
</tr>
</tbody>
</table>

Practice

We will now rate several responses together. After reading or listening to each response, please write down the score (number from 1 to 4) that you would assign to the response.

Sample Item A rating: ______
Sample Item B rating: ______
Sample Item C rating: ______
Sample Item D rating: ______
Sample Item E rating: ______
Sample Item F rating: ______
Rating of spoken discourse completion tasks. Refer to the ODCT handouts to remind yourself of the prompts as needed. Please rate each response from 1 to 4.

**Disc Number __________**

**Person Number one:**

<table>
<thead>
<tr>
<th>Response 1</th>
<th>Response 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 2</td>
<td>Response 8</td>
</tr>
<tr>
<td>Response 3</td>
<td>Response 9</td>
</tr>
<tr>
<td>Response 4</td>
<td>Response 10</td>
</tr>
<tr>
<td>Response 5</td>
<td>Response 11</td>
</tr>
<tr>
<td>Response 6</td>
<td>Response 12</td>
</tr>
</tbody>
</table>

**Person Number two:**

<table>
<thead>
<tr>
<th>Response 1</th>
<th>Response 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 2</td>
<td>Response 8</td>
</tr>
<tr>
<td>Response 3</td>
<td>Response 9</td>
</tr>
<tr>
<td>Response 4</td>
<td>Response 10</td>
</tr>
<tr>
<td>Response 5</td>
<td>Response 11</td>
</tr>
<tr>
<td>Response 6</td>
<td>Response 12</td>
</tr>
</tbody>
</table>

**Person Number three:**

<table>
<thead>
<tr>
<th>Response 1</th>
<th>Response 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response 2</td>
<td>Response 8</td>
</tr>
<tr>
<td>Response 3</td>
<td>Response 9</td>
</tr>
<tr>
<td>Response 4</td>
<td>Response 10</td>
</tr>
<tr>
<td>Response 5</td>
<td>Response 11</td>
</tr>
<tr>
<td>Response 6</td>
<td>Response 12</td>
</tr>
</tbody>
</table>
APPENDIX I

TRANSCRIPT OF PRAGMATIC VIGNETTES

Homework Request

R: Hey, what did he say about the homework assignment?
M: He said we have an extra day to finish and turn it in.
R: Okay. Um, have you done yours yet?
M: I’ve started it, but, um, I’d have to work on it a couple of hours to get it done.
R: Alright, um, do you think I could take a look at yours?
M: Um. The thing is, I haven’t printed it out yet, and I’m not planning to work on it until later on tonight.
R: Okay, well that’s cool. Um, how about we just talk about it tomorrow?
M: Sounds good!
R: Alright!

Classroom Request

K: Okay everyone, would you please open your books to page 37? Let’s do the questions on page 37, please.
D: Dr. Kott?
K: Deborah, yes.
D: Yes, I was wondering, could we review the old questions before we start the new questions?
K: Oh, well, that would be a very good idea, but I think I gave everyone two days to finish that assignment, and you may have finished, but the others perhaps are still working through it, so maybe it would be better to wait until tomorrow for that.
D: I had just, I forgot you gave us an extra day.
K: Oh, well, please, Deborah, please, could you begin on page 37 for us then?
D: Sure.
K: Thank you.

Classroom Request (2)

K: Okay everyone, that’s it for this week. Thank you very much. Have a nice weekend. I’ll see everybody on Monday. Thank you so much!
K: Uh, Molly, please, yes.
M: Dr. Kott,
K: Yes.
M: I'm not really clear. Could you repeat that for me?
K: Well, what I was talking about during the lecture is that we have to keep in mind that, in order to finish the project, there are those principles that you have to keep in mind and go over.
M: Yeah, I wasn't really clear about those. Do you think you could, I don’t know, maybe put them on the website?
K: Uh, that's a good idea. I would like to very much, but unfortunately, I don’t have those—for
the moment, I don’t have them on electronic format.
M: Oh, cool, I mean, I just thought I’d ask.
K: OK. Thank you very much. Have a nice weekend.
M: Bye, have a nice weekend.
K: Thank you.

Ride Request

R: Hey, Justin.
J: Hey, Ramin!
R: How you doin’?
J: Good, good. How have you been?
R: Good, man. I can’t complain.
J: Alright.
R: Yeah.
J: Hey, I was going to ask you, uh, are you going to be going to that big basketball game this
weekend?
R: Definitely.
J: Cool.
R: Yeah.
J: Uh, do you think I could get a ride with you, maybe? I’m uh, I’m planning to go, but…
R: Man, I wish I could, but I’m going to be meeting up with some friends earlier, and I don’t
know where we’re going to be leaving from, so…
J: OK. Hey, that’s alright. I understand. I’m going to try to make it there anyway.
R: OK.
J: Alright.
R: Hey, sorry about that.
J: Alright. We’ll see you.

Conference Request

D: Hi, Dr. Kennell.
P: Hi, Deborah.
D: Yeah, uh, you had asked for some students to come by to help you with the class project?
P: Right, yeah. I’m hosting a conference next month and I really need some students to help
with the logistics of it.
D: Yeah? What kinds of things do you need help with?
P: Well, I’ve got some students working on the fliers and the, uh, programs, and I just need
someone to help with the posters.
D: Posters? I’m not really much of an artist.
P: That’s OK. There are plenty of other things we can do that you can help with.

Extension Request

J: Hey, Dr. Kennell?
P: Yes.
J: Yeah, um, I was going to tell you, the class is great. I'm really having—I feel like I'm learning a lot. The thing is on the paper, um, I was wondering if it’d be possible to get an extension. I need, maybe, until Thursday. Is that—Is that going to be possible?
P: It’s not—I wish I could help you out, but it’s not going to be possible. Um, deadline is firm. Uh, I’m going to be out of town Wednesday.
J: Oh, I see.
P: And, uh, I just can’t, I just can’t offer you a deadline—I mean, uh, an extension on this.
P: OK.
J: Thank you.
P: You’re welcome.

Car Advice

R: Hi there. Can I help you?
L: Hi, um, yeah, my brakes have been making this really loud, funny noise lately. And I was wondering if you could look at them for me.
R: Sure, let’s see if we can see anything…Yep, looks like your brake pads are worn down there. You really should change those out as soon as possible.
L: How much is that going to cost?
R: Probably about forty dollars for each pair.
L: You mean I have to change both sides?
R: You really should go ahead and take care of both sides now, so down the road you don’t have uneven wear.
L: Gosh. Okay, go ahead and change both pairs then.
R: Alright, sounds good. We’ll take care of it for you, alright?
L: Okay. Thank you.
R: You’re welcome.

Locker Room Advice

R: Hey, how’s it going?
J: Pretty good. How was the game?
R: It was good. It was good. You know, my knee’s been hurting me a little bit, though.
J: Oh man, does that happen every time you play?
R: Uh, mostly when I'm running and jumping.
J: OK. Well, why don’t you—you ought to try wrapping it before the game.
R: Yeah, that’s what Victor says.
J: Well, he’s had enough knee problems of his own. He ought to know. Maybe you should listen to him.
R: That’s true. That’s true. The thing is, I don’t like how it makes me feel. I feel like I’m being constricted when I run and jump when it’s on my knee. I don’t know.
J: Well, you could try wrapping it kind of loosely at first, and then loosen or adjust it as the game goes on.
R: Right, OK. It could work. It could work.
J: Alright. Give it a shot.
R: Alright.

Café Advice

R: So, Laura, how’s everything at school?
L: Oh, everything is great. I mean, it’s really difficult. My classes are demanding. But I’ve met a lot of new people. It’s a lot of fun.
R: Oh, that’s great. You must have something on your mind, because you invited me to lunch. Do you have anything to talk about?
L: Well, yes. I really just wanted to ask your advice because I can’t decide if it’s a crazy plan or not to spend a semester in Europe. I’m really not sure exactly what I should do.
R: A semester in Europe?
L: Mm-hmm.
R: Well, are you planning to come back to school?
L: Well of course! I mean, school’s the most important thing in my life.
R: Well, if you want to go to Europe, you should do it now while you’re young enough to appreciate it. And I’m sure if you travel around Europe and visit all the capital cities, it’s going to be just as educational as staying here for a year. So, I think you should do it now.
L: Oh thanks, Dad. I feel so much better talking to you about my decision.
R: Well, please. Listen, we raised you to have a wonderful, good head on your shoulders. So I think you should just go and enjoy yourself and have fun. And now that I’ve helped you, is there anything you can suggest for lunch?
L: Well, you could try the coconut shrimp here. They’re really good.
R: Yes? Do you think I should order that?
L: Yes.
R: OK. Thank you very much. It’s so nice being with my daughter.
L: Thanks, Dad!

Test Compliment

J: Hey man, why weren’t you at the game Saturday?
R: Oh man, I had too much studying to do. Professor K gave us a big test yesterday.
J: Oh, man. Well, how’d you do? I hope all that studying paid off; you missed a great game!
R: Well, I got a 96, so I can’t complain.
J: 96. Man, that's awesome! I wish I could be that good of a student.
R: Thanks man, it wasn't easy.
R: Well, I just really wanted to do well on this one.
R: Thanks, but you’re doing pretty well in this class, even without all the cramming.
R: Well, I was actually pretty surprised.

Haircut Compliment

D: Hey!
M: Hey!
D: How was your weekend?
M: Oh, it was good. I just got to relax a little bit.
D: Yeah?
M: Yeah.
D: That's great. Did you get your hair cut?
M: Yeah, I did.
D: Yeah. I think it looks great!

M: Thanks!
M: Thanks. I feel like they cut it a little too short, though, I don’t know, what do you think?
M: Oh, yeah, well, I wanted to do something a little different with it. I don’t know.
M: Oh thanks. You’re so sweet to say so!

Paper Compliment

R: Hey Dr. Kennell, how you doin’?
P: Hey Ramin, how are you doing?
R: Good, thanks. I just wanted to talk to you about the paper for our class.
P: Yeah, did you bring it?
R: I did. This is it, here.
P: Let me take a look at it real quick. [reading] Yeah. This conclusion is really—it’s much better, very strong. I like how you explained it here.
P: Great.
P: Good job.

R: Oh, thank you. I wasn't sure if I had included enough support.
R: Oh well, I just followed the example that you gave us in class.
R: Hey, I stayed up all night to get it done, but it was worth it.
R: Thank you. I’m happy to hear that.

Lunch Invitation

L: Hey, how’s it going?
K: Hey, how have you been?
L: Pretty good, but I’ve been so busy with my new teaching job, I haven’t been keeping up with anything lately.
K: Oh, I understand.
L: Um, how’s things been going with you?
K: Oh, I can’t complain. What brings you to this neck of the woods?
L: Oh, well, I was just in the library, but now I have to go to the County Clerk’s office.
K: Well, hey, do you want to grab a bite to eat later?
L: That’d be great!
K: Mm, how about eight o’clock?
L: That sounds good.
K: Okay, great.
Book Signing Invitation

J: Now, who’s this to?
L: Uh, to Laura.
J: Well, thank you for coming out to the book signing today.
L: Oh, it’s really a privilege to be here.
J: Did you know Dr. Jenks?
L: Yes, I was a student of Dr. Jenks several times, up until his retirement last year.
J: Well, because you know him, you might be interested. We’re going to have a formal
    ceremony tomorrow honoring Dr. Jenks.
L: Oh, really?
J: Yes, so I could extend an invitation to you. Dr. Jenks will be there.
L: Oh, thank you so much for the invitation. I’ll plan on being there.
J: Not at all. Now, it’s semi-formal, but you don’t need to bring anything. So just talk to my
    secretary there and he’ll give you all the details.
L: Thanks again so much for the invitation.
J: You’re quite welcome.

Party Invitation

R: Joe, how’s it going?
J: Pretty good. How have you been?
R: Good, good. What’s going on?
J: Not much. I’m just busy. You know, I’ve got a lot of stuff going on with school.
R: Listen, I’m glad I ran into you. We’re having a party this Friday.
J: Oh, really?
R: Yeah. You should come over.
J: I mean, I’d like to, but I don’t know. I’m really busy with school and that kind of thing.
R: Uh-huh. Well hey, listen man, all work and no play. Come on!
J: Yeah.
R: Take a break. Come to the party.
J: I could use a little time to unwind. There’s always tomorrow!
R: Absolutely.
J: Alright.
R: And hey, you can bring your roommate, too, alright? If he’s recovered from the last party.
J: Yeah, you know him.
R: Alright. I’ll see you then.
J: I appreciate it.

Presentation Invitation

T: Hi, Dr. Smith. Good to see you again.
K: Terry, how are you? Nice to see you.
T: Good, good.
K: Well, this research paper looks really good.
T: Oh, I appreciate that. I worked pretty hard on it, so…
K: Well, it’s important to let the students be recognized for good work.
T: OK. Well, thanks, yeah.
K: Uh, would you be interested in presenting your findings at the annual conference next week?
T: Oh, at the conference? Oh. I mean, I’d like to, but I’m not sure I’m ready for something that big.
K: Well, your professors and I would be happy to work with you on how to present it and how to present your findings and answer any questions.
T: OK, alright. I think I’ll probably do it then. Let me get back to you, but it sounds really good.
K: Great! We’ll set up an appointment, a meeting for next week.
T: Ok, we’ll do that. Thank you.
T: Thank you.
APPENDIX J

COMPREHENSION AND EXTENSION QUESTIONS FOR – OUTPUT INSTRUCTION

Classroom Request
1. Which page is the class working on?
   a. 200    b. 73    c. 10    d. 37
2. How many days did the professor give the class to finish the old assignment?
   a. one    b. two    c. three    d. four
3. When does the professor suggest they wait until to go over the old questions?
   a. “later today”    b. “next week”    c. “tomorrow”    d. it is not stated in the video
4. Why did she forget that they had been given an extra day?
   a. she has a poor memory    b. she was absent from class    c. it is not stated in the video
5. Do you prefer to work on homework problems on your own or with your classmates?
   a. on my own    b. with classmates
6. Why did you make that choice?

Test Compliment
1. What day was the big game?
   a. Sunday    b. Saturday    c. Monday    d. it is not stated in the video
2. Why did the second student stay at home?
   a. he was tired    b. he forgot about the game    c. he had a big test
3. What score did he receive on the big test?
   a. 96    b. 92    c. 69    d. 100
4. What was the name of the professor who gave the test?
   a. Professor J.    b. Professor A.    c. Professor K.    d. it is not stated in the video
5. Do you get anxious (nervous) when you have a big test?
   a. yes    b. no    c. I’m not sure
6. Why do you think you do (or do not) get anxious?
**Classroom Request (2)**

1. On what day do you think this scene takes place?
   a. Monday   b. Friday   c. Wednesday   d. there is no way to know
2. How do you know?
3. What was the lecture about?
   a. some principles   b. a class project   c. both a and b   d. none of these
4. Where does she want the professor to post the notes?
   a. on the wall of the classroom   b. on television   c. on the course website
      d. it is not stated in the video
5. Do you think using the internet to post class notes is a good idea?
   a. yes   b. no   c. I’m not sure
6. Why or why not?

**Haircut Compliment**

1. What did the second woman do over the weekend?
   a. traveled   b. worked on homework   c. relaxed   d. visited the Capitol building
2. Did she get her hair cut?
   a. yes   b. no   c. it is not stated in the video
3. Where did she get her hair cut?
   a. at the salon   b. at a new hairstylist   c. it is not stated in the video
4. How often do you get your hair cut?
   a. once a month   b. once a week   c. every two weeks   d. other
5. How do you feel when you get a new haircut?
   a. confident   b. happy   c. worried   d. I’m not sure / other
6. Why do you feel that way?

**Ride Request**

1. What kind of game is Ramin going to?
   a. football   b. basketball   c. baseball   d. soccer
2. When is the game taking place?
   a. the same day   b. next week   c. on the weekend   d. it is not stated in the video
3. What does Justin want from Ramin?
   a. a ride to the game   b. to borrow some money   c. a ticket to the game
   d. notes from class
4. Why can’t Ramin help him?
   a. he has no money   b. he has no time   c. he’s meeting some friends
   d. it is not clear in the video
5. Does Justin still plan to try to go to the game?
   a. yes   b. no   c. it is not clear from the video
6. What kinds of sports do you like to play or watch?

**Paper Compliment**
1. What did Ramin write the paper for?
   a. for a conference   b. for Dr. Kennell’s class   c. for a job application
   d. it is not stated in the video
2. How many pages long is the paper?
   a. 20 pages   b. 15 pages   c. 10 pages   d. it is not stated in the video
3. Did Ramin bring the paper with him?
   a. yes   b. no   c. it is not clear in the video
4. Which section of the paper does Dr. Kennell comment about?
   a. the introduction   b. the body   c. the conclusion   d. it is not stated in the video
5. Are you good at writing papers in English?
6. Are you a good writer in your first language?

**Conference Request**
1. Why did Deborah come by to see Dr. Kennell?
2. When is the conference?
   a. next week   b. next month   c. in two months   d. it is not stated
3. Who is hosting the conference?
   a. Dr. Kennell   b. Dr. Smith   c. Deborah   d. it is not clear in the video
4. Which of these things are students already helping with?
   a. fliers   b. posters   c. programs   d. both (a) and (c)
5. Why doesn’t Deborah want to help with the posters?
   a. she does not like posters  b. she has no time  c. she is not a good artist 
   d. the reason is not clear in the video
6. Do you consider yourself to be creative or artistic? Why or why not?

Extension Request
1. Does the student enjoy Dr. Kennell’s class?
2. Until what day does Justin want to extend the deadline for the paper?
   a. Tuesday  b. Wednesday  c. Thursday  d. it is not stated in the video
3. When is Dr. Kennell going to be out of town?
   a. Tuesday  b. Wednesday  c. Thursday  d. Friday
4. Where is he going?
   a. New York  b. to a conference  c. to Canada  d. it is not stated in the video
5. Is he willing to extend the deadline for the paper?
6. Have you ever missed the deadline for a paper or another assignment? What was the result?

Bank Advice
1. Where is this dialog taking place?
   a. a hotel  b. a bank  c. a school  d. a kitchen
2. Will the man withdraw his money more than three times a year?
   a. yes  b. no  c. it is not clear
3. What type of account does the woman recommend?
   a. checking account  b. money market savings account  c. certificate of deposit
   d. regular savings account
4. What are the two benefits of the account that she recommends?
   a. _______________________________________________________________________
   b. _______________________________________________________________________
5. What does the man ask the woman about the type of account that she recommends?
   a. if it is easy to access  b. if it is riskier than a regular account  c. if it is new
   d. if it is as good as a regular account
6. Do you think it is important to save money for the future? Why or why not?

Lunch Invitation
1. What is the first woman’s new job?
   a. lawyer   b. secretary   c. dentist   d. teacher
2. What is the meaning of the expression “this neck of the woods”?
   a. far away   b. this week   c. this area   d. anyplace
3. What two places does the first woman need to go?
   a. __________________________   b. __________________________
4. What do you think is the function of the County Clerk’s office?
5. What time do the two women agree to meet to eat?
   a. 8 o’clock   b. 7 o’clock   c. 6 o’clock   d. it is not stated in the video
6. At what times do you usually eat lunch and dinner?

Car Advice
1. What part of the woman’s car is giving her a problem?
   a. the steering wheel   b. the brakes   c. the tires   d. the wipers
2. What does she say is the problem?
   a. there is a loud noise   b. they don’t work   c. they are missing
   d. she does not say
3. What does the mechanic say is the problem?
   a. the pads are missing   b. the pads are old   c. the pads are worn down
   d. the pads are rough
4. How much will it cost to change both pairs?
   a. $20   b. $100   c. $40   d. there is no charge (free)
5. Does the woman agree to have them changed?
6. Have you had an experience at the car repair shop? Briefly describe your experience.

Book Signing Invitation
1. Who is the book about?
   a. Dr. Kennell   b. Dr. Jenks   c. Dr. Smith   d. it is not clear
2. How does Laura know him?
   a. she was his secretary   b. she read a book about him   c. she met him at a meeting
       d. she was a student of his
3. When did he retire?
   a. last month   b. last week   c. last year   d. he is not retired yet
4. When will the formal ceremony be?
   a. tomorrow   b. next week   c. next month   d. it is not stated in the video
5. What kind of event is the ceremony?
   a. formal   b. semi-formal   c. casual   d. it is not stated in the video
6. Have you ever attended a formal or semi-formal event? What was the event?

**Locker Room Advice**
1. What kind of game was Ramin playing in?
   a. soccer   b. football   c. basketball   d. it is not stated in the video
2. When does Ramin’s knee usually hurt him?
   a. when he runs   b. when he walks   c. when he jumps   d. both (a) and (c)
3. What does Justin suggest for Ramin to do?
   a. rub his knee   b. stop playing   c. wrap his knee   d. he does not give any advice
4. Who is the other person who says it’s a good idea to wrap the knee?
5. Why does Ramin not want to follow the suggestion?
   a. it feels constricting   b. it is expensive   c. it is too difficult to do
       d. it is not stated in the video
6. Have you ever had an injury? What kind of injury was it, and how did you deal with it?

**Party Invitation**
1. What is keeping Joe so busy?
   a. work   b. exercising   c. school   d. it is not clear from the video
2. When is Ramin’s party?
   a. Saturday   b. Friday   c. next weekend   d. it is not stated in the video
3. What do you think “all work and no play” means?
4. What expression does Joe use in response?
   d. “Tomorrow’s another day”
5. Who else is invited to the party?
   a. Joe’s best friend   b. Joe’s roommate   c. Joe’s brother   d. it is not clear
6. Describe a great party that you have attended. What made it memorable?

**Café Advice**
1. Where is this scene taking place?
   a. a party   b. a classroom   c. a café   d. Europe
2. Where is Laura thinking about going?
   a. Japan   b. Europe   c. Mexico   d. it is not stated in the video
3. For how long does she plan to go?
   a. two weeks   b. a month   c. a session   d. a semester
4. Does Laura plan to return to school?
   a. Yes   b. No   c. It is not clear in the video
5. Does her father (Roger) agree with the idea?
   a. Yes   b. No   c. It is not clear in the video
6. Do you think Laura’s idea is a good one?

**Presentation Invitation**
1. What does Dr. Smith think about Terry’s research paper?
   a. it is too short   b. it is very good   c. it is too confusing   d. it is not stated in the video
2. Where does Dr. Smith want Terry to present his research paper?
   a. at a lunch meeting   b. in a class   c. in his office   d. at a conference
3. Who will work with Dr. Smith to help Terry prepare?
   a. Terry’s professors   b. Terry’s classmates   c. Terry’s parents   d. it is not clear
4. Does Terry agree to present his paper?
   a. Yes   b. No   c. Probably   d. We do not know from the video
5. When does Dr. Smith want to meet with Terry?
   a. next month   b. next week   c. tomorrow   d. next year
6. Do you like to give presentations in front of other people?

**Presentation Compliment**

1. What is the topic of Molly’s presentation?
   a. Japanese food  
   b. Korean food  
   c. Chinese food  
   d. it is not stated in the video

2. Does Ramin like the presentation?
   a. Yes  
   b. No  
   c. it is not clear from the video

3. Did Molly work hard to prepare the presentation?
   a. Yes  
   b. No  
   c. it is not clear from the video

4. Was she nervous?
   a. Yes, very nervous  
   b. Yes, a little  
   c. No  
   d. it is not clear in the video

5. Where is this dialogue taking place?
   a. a restaurant  
   b. a business meeting  
   c. a classroom  
   d. a kitchen

6. Which would you prefer, American food or Korean food?
APPENDIX K

HUMAN SUBJECTS APPROVAL MEMORANDUM

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673 • FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 2/23/2006

To: Justin Jernigan
MC 4490

Dept.: MIDDLE AND SECONDARY EDUCATION

From: Thomas L. Jacobson, Chair

Re: Use of Human Subjects in Research
Instruction and Developing Second Language Pragmatic Competence: An Investigation into the Efficacy of Output

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Exempt per 45 CFR § 46.101(b) 2 and has been approved by an accelerated review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If the project has not been completed by 2/22/2007 you must request renewed approval for continuation of the project.

You are advised that any change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must promptly report in writing, any unexpected problems causing risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols of such investigations as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Protection from Research Risks. The Assurance Number is IRB00000446.

Cc: Eleni Pappamichail
HSC No. 2006.0173
APPENDIX L

INFORMED CONSENT FORM

I freely and voluntarily and without element of force or coercion, consent to be a participant in the research project entitled “Instruction and Developing Second Language Pragmatic Competence: An Investigation into the Efficacy of Output.”

This research is being conducted by Justin E. Jernigan, M.A., a doctoral student in the Department of Middle and Secondary Education, Multilingual/Multicultural Education Program at Florida State University, working under the supervision of Dr. N. Eleni Pappamihili. The purpose of this research project is to better understand the ways in which classroom instruction may facilitate the acquisition of pragmatic skills in English among adult students of English as a second language. I understand that if I participate in the project I may be recorded by the researcher. The purpose of the recordings is to allow the researcher to better analyze and describe the learning process. These tapes will be kept by the researcher in a locked file drawer. I understand that only the researcher will have access to these tapes and that they will be destroyed by December 31, 2007.

I understand I will also be asked to provide samples of my written and spoken English, and that I will be assessed as to my ability to correctly produce some difficult grammatical forms. I agree to allow the researcher to keep my written documents for the purposes of analysis and identifying connections between classroom behavior and learning of English. These written documents will also be destroyed by December 31, 2007.

I understand my participation is totally voluntary and I may stop participation at any time. If I decide to stop participation, I will still be entitled to view the summary of final study results. Information obtained during the course of the study will remain confidential to the extent allowed by law, and will be identified by a subject code number. My name will not appear on any of the results.

I understand that there is minimal risk involved if I agree to participate in this study. No conditions are anticipated that fall outside of the scope of normal classroom study and interaction. I am also able to stop my participation at any time I wish.

I understand there are benefits for participating in this research project. My own awareness about my learning may be increased. Also, I will be providing researchers studying second language acquisition with valuable insight into how classroom interaction affects the acquisition of pragmatic skills in English. This knowledge could help teachers in second language classrooms provide better instruction to their students.

I understand that my consent may be withdrawn at any time without prejudice, penalty or loss of benefits to which I am otherwise entitled. I have been given the right to ask and have answered any inquiry concerning the study. Questions, if any, have been answered to my satisfaction. A summary of the study results will be given to me upon my request.

I understand that I may contact Justin E. Jernigan, Florida State University, Dept. of Middle and Secondary Education, 209 Carrothers Hall, Tallahassee, FL 32306-4490, (850) 644-6553, jej03@garnet.acns.fsu.edu, for answers to questions about this research or my rights.

I may also contact Mr. Jernigan’s supervising professor, Dr. N. Eleni Pappamihili, Florida State University, Dept. of Middle and Secondary Education, 209 Carrothers Hall, Tallahassee, FL 32306 4490, (850) 644-6553, pappamih@coe.fsu.edu.

If you have any questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk, you can contact the Chair of the Human Subjects Committee, Institutional Review Board, through the Vice President for the Office of Research at (850) 644-8633.

I have read and understand this consent form.

(Subject)  (Date)
REFERENCES


Quarterly, 27, 33-56.


through classroom interaction (pp. 1-20). Mahwah, NJ: Lawrence Erlbaum Associates.


Blackwell.


183


Yuan, Y. (2001). An inquiry into empirical pragmatics data-gathering methods: Written DCTs,

BIOGRAPHICAL SKETCH

Justin E. Jernigan is the son of Glenn L. Jernigan and Nancy J. Garner, both of Live Oak, Florida. He obtained a B.A. in Russian (minor in linguistics) in 1994 and M.A. in Linguistics with TESL certification in 1997, both from the University of Florida. He also earned the M.Div. from the Southern Baptist Theological Seminary in 2003. Justin has lived for significant periods in Live Oak, Florida; Gainesville, Florida; Louisville, Kentucky; and Northwest Malaysia. He currently lives in Tallahassee, Florida with his wife, Vien, and two sons. In Tallahassee, Justin has been able to pursue the Ph.D. degree in Multilingual/Multicultural Education and grow as a teacher at the Center for Intensive English Studies and as a graduate teaching assistant in the Department of Middle and Secondary Education at Florida State University.