FACTORS INFLUENCING ORAL CORRECTIVE FEEDBACK PROVISION IN THE SPANISH FOREIGN LANGUAGE CLASSROOM: INVESTIGATING INSTRUCTOR NATIVE/NONNATIVE SPEAKER STATUS, SLA EDUCATION, & TEACHING EXPERIENCE

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in the Department of Spanish and Portuguese

By

Laura Gurzynski-Weiss, M.S.

Washington, DC
May 25, 2010
ACKNOWLEDGEMENTS

There are so many to thank for their support and guidance throughout this process. It is rare to find world-renowned scholars who are also wonderful people who encourage a balance between work and life; I was blessed to work with several. To my mentor, the one and only Ron Leow: thank you for your unending support and for the opportunities you have given me throughout the past five years. Thanks for letting me make my own decisions and mistakes along the way, and for listening to my rationale and always giving me more to think about. I will always be grateful to have “Research Ron” on my side as a mentor and friend. To my two incomparable committee members, Alison Mackey and Andrea Révész: Alison, there aren’t many who can say they were introduced to interaction-based research by Alison Mackey and I am truly grateful for having had the opportunity. Thank you for being the voice of reason when I wanted to do too much, for the incredible opportunities to work with you as an RA, and for your faith in the importance of instructor-focused classroom research from the beginning. Andrea, thank you for introducing me to task-based language teaching and research and for encouraging me to bridge the gap between my research and TBLT. You have been the ideal guide for the end-of-graduate school journey and I am honored to have had the chance to work and present together.

To Rusan Chen, statistician extraordinaire and one of the most positive people I have ever met: thank you for your invaluable assistance with statistical analyses and for always being such a motivating force in my graduate school career. To the Spanish FL instructors who participated in this study: your enthusiasm, generosity of time and
support of the project were truly an inspiration both professionally and personally. Thank you to Susan Pennestri for always lending me equipment and saving me from tech crises that arose during data collection. A large thank you is due to the professors and grad students I worked with: this final product is the culmination of years of experiences and challenges from working with each and every one of you. Thank you to Almitra Medina and Nina Moreno, who offered advice from someone who has “been there,” and to Ellen Johnson and Julio Torres for helping me with errands while I finished writing from WI.

To the “fam” I was blessed with at GU: thank you for the countless hours of conversation and camaraderie the last five years. To Michelle Farrell, who was always willing to go for walks and out dancing, no matter how much we had to do. To Grant Armstrong: the way you are willing to share your genius with those around you is incredible. Thanks for explaining ‘trees’ to me and for having many profound talks over drinks. To Melissa Baralt, my BFF in projects and all grad school adventures: your friendship made a difference so many times and I am truly grateful for it and for you.

To my parents, who always believed in me every step of this journey and gave me the confidence and encouragement to pursue my dreams, no matter how far away they took me. To my dad, who is always proud of me no matter how small the success: you always make me feel so special and loved. Thank you for always being there for me and for having faith in my decisions. To my mother, a virtual lifeline on the phone and my layperson consultant for countless ideas: thank you for always making the time to talk with me, no matter how long the conversation or how many calls a day you got from me. Thank you especially for being both my mom and best friend. I am truly blessed.
Lastly, to my husband, Nick, for putting up with my “grad school affair” for the last five years of our relationship and the first two of our marriage: thank you for always giving me a healthy dose of perspective and laughter and for forcing me to be human and “step away from the computer!” Thank you for your understanding, your sense of humor, and for getting excited with me about findings when you didn’t have a clue what I was talking about. You are an amazing partner in love and life.
FACTORS INFLUENCING ORAL CORRECTIVE FEEDBACK PROVISION IN THE SPANISH FOREIGN LANGUAGE CLASSROOM: INVESTIGATING INSTRUCTOR NATIVE/NONNATIVE SPEAKER STATUS, SLA EDUCATION & TEACHING EXPERIENCE

LAURA GURZYNISKI-WEISS, M.S.

THESIS ADVISOR: RONALD P. LEOW, PH.D.

ABSTRACT

The role of interactional feedback has been a critical area of second language acquisition (SLA) research for decades and while findings suggest interactional feedback can facilitate SLA, the extent of its influence can vary depending on a number of factors, including the native language of those involved in communication. Although studies have found differences in the ways native (NS) and non-native speakers (NNSs) provide feedback, most have compared non-teaching individuals outside of the classroom; the few comparing NS and NNS language instructors have been largely limited to the English as a Second Language (ESL) setting, have typically measured teaching differences indirectly via questionnaires, and focused on topics other than feedback. Two studies have begun to explore additional instructor individual difference (ID) factors, education and experience, in relation to feedback provision in ESL learning contexts (Mackey, Polio, & McDonough, 2004; Polio, Gass, & Chapin, 2006). However, to date there has been no thorough examination of instructor ID factors in relation to naturally occurring feedback in the foreign language (FL) classroom.

The current study sought to amplify this minimal research by investigating three instructor ID factors — NS/NNS status, SLA education, and years of teaching experience — in the same study. Using quantitative and qualitative data gathered from 60 Spanish
FL instructors, the study investigated (a) the amount and type of feedback instructors provide during a 50-minute lesson, (b) if they differ regarding the factors they take into account while making in-class feedback decisions, and (c) if these differences are related to the aforementioned ID factors. This dissertation also examined whether instructors’ feedback beliefs correspond with their in-class provision, and whether there were belief differences relating to instructor NS/NNS status, SLA education or teaching experience.

Analyses revealed that in-class feedback provision was significantly related to each of the ID factors. However despite these behavioral differences instructors did not differ in their beliefs; thus, there were few significant relationships between beliefs and corresponding practices. Finally, data from 35 stimulated recall protocols shed light on the nature of these differences and how instructor ID factors influence their beliefs and in-class feedback decisions.
# Table of Contents

List of Appendices ........................................................................................................ xvi

List of Figures ................................................................................................................ xvii

List of Tables .................................................................................................................. xviii

**CHAPTER ONE: Introduction** ................................................................. 1

  Statement of the Problem ................................................................. 1

  Definition of Terms ........................................................................... 6

    Beliefs ......................................................................................... 6

    Corrective Feedback/Error Correction/Negative Feedback .......... 7

    Explicit Feedback ....................................................................... 7

    Implicit Feedback ........................................................................ 7

    Individual Differences/Factors .................................................. 7

    Interlocutor .............................................................................. 7

    Native Speaker .......................................................................... 7

    Oral Corrective Feedback ......................................................... 8

    Practice ..................................................................................... 8

    Stimulated Recall Protocol ......................................................... 8

**CHAPTER TWO: Review of the Literature** ........................................... 9

  Theoretical Background ............................................................... 9

    The Interaction Hypothesis ....................................................... 9

    The Output Hypothesis ............................................................. 10

    Schmidt’s Noticing Hypothesis .................................................. 10
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>11</td>
</tr>
<tr>
<td>CORRECTIVE INTERACTION-BASED FEEDBACK</td>
<td>12</td>
</tr>
<tr>
<td>Overview</td>
<td>12</td>
</tr>
<tr>
<td>Internal Factors</td>
<td>14</td>
</tr>
<tr>
<td>Learner Proficiency Level and Developmental Readiness</td>
<td>14</td>
</tr>
<tr>
<td>Learner Age</td>
<td>15</td>
</tr>
<tr>
<td>Learner Anxiety</td>
<td>16</td>
</tr>
<tr>
<td>Learner Motivation and Attitudes Towards Feedback</td>
<td>17</td>
</tr>
<tr>
<td>Learner Aptitude and Working Memory</td>
<td>17</td>
</tr>
<tr>
<td>External Factors</td>
<td>19</td>
</tr>
<tr>
<td>Degree of Explicitness and Feedback Type</td>
<td>19</td>
</tr>
<tr>
<td>Amount of Feedback</td>
<td>21</td>
</tr>
<tr>
<td>Setting of Feedback Provision</td>
<td>23</td>
</tr>
<tr>
<td>Mode of Feedback Provision</td>
<td>25</td>
</tr>
<tr>
<td>Length of Intervention</td>
<td>26</td>
</tr>
<tr>
<td>Feedback Target</td>
<td>27</td>
</tr>
<tr>
<td>Feedback Focus</td>
<td>28</td>
</tr>
<tr>
<td>Task-Related Factors</td>
<td>28</td>
</tr>
<tr>
<td>Feedback Provision According to Dyad</td>
<td>29</td>
</tr>
<tr>
<td>Feedback Provision According to Recipient</td>
<td>32</td>
</tr>
<tr>
<td>Source of Feedback Provision</td>
<td>33</td>
</tr>
<tr>
<td>Summary of Interaction-Based Feedback Research</td>
<td>35</td>
</tr>
</tbody>
</table>
INSTRUCTOR NATIVE/NON-NATIVE SPEAKER STATUS ……….. 36

Overview………………………………………………………… 36

Comparing Self-Perceptions of NS and NNS Instructors ………40

Student Perceptions of Instructor NS/NNS Status ……………43

NS and NNS Instructors’ Intuitions in the Target Language ……45

NS and NNS Instructors’ Teaching Behavior ……………………48

  Error Evaluation Outside of the Classroom ………………….49

  Feedback Provision in the Classroom ……………………….. 51

Strengths/Weaknesses of Existing NS/NNS Instructor Research … 53

INSTRUCTOR EDUCATION ……………………………………… 61

Overview………………………………………………………… 61

Existing Research on Instructor Education ……………………. 62

Limitations of Existing Research/ Motivations for Investigating
  Instructor SLA Education …………………………………….. 64

INSTRUCTOR TEACHING EXPERIENCE …………………….. 66

Overview………………………………………………………… 66

Operationalizing Instructor Experience ……………………... 71

Comparing More and Less Experienced Instructors …………. 75

Limitations of Existing Instructor Experience Research ……… 80

Comparing Instructor Beliefs and Corresponding Practices ……… 82

Overview ………………………………………………………… 82

Research Investigating Beliefs About Corrective Feedback …… 83
Studies Comparing Feedback Beliefs and In-Class Provision …..83

Limitations to Existing Belief/Practice Research on Feedback …..86

Rationale for the Current Study ..................................................87

Research Questions .................................................................90

Hypotheses .............................................................................91

CHAPTER THREE:  Research Design and Methodology ................. 93

Introduction .............................................................................93

Operationalizations ...............................................................93

Native/Nonnative Speaker Status ..........................................93

SLA Education ..................................................................94

Teaching Experience ..........................................................94

Description of Participants and Independent Variables .......... 96

Research Design .................................................................100

Procedure ........................................................................... 101

Participant Recruitment .....................................................102

Recordings of Spanish Language Classes ............................103

Stimulated Recall Protocols ..................................................107

Stimulated Recall Methodological Considerations ..............109

Questionnaire Dissemination .................................................111

Materials .............................................................................111

Participant Background Questionnaire .................................111

Stimulated Recall Protocol Questions .................................113
Participant Belief Questionnaire ........................................... 114
Pilot Studies ........................................................................... 1116
Overview ............................................................................... 116
Pilot Participants ................................................................. 119
Results: Participant Background Questionnaire .................. 120
Results: Recordings of Spanish Language Classes .............. 121
Results: Stimulated Recall Protocols ................................. 127
Results: Participant Belief Questionnaire ......................... 134
Conclusions/Changes Resulting from the Pilot Studies ....... 137
Coding Procedures ................................................................. 137
Participant Background Questionnaire ......................... 138
Classroom Interaction Data ................................................. 139
Stimulated Recall Protocols ................................................. 145
Participant Belief Questionnaire ................................. 146
Statistical Procedures ............................................................... 146
Research Question 1: Instructor Beliefs ......................... 147
Research Question 2: Instructor In-Class Feedback Provision .... 147
Research Question 3: Consistency of Feedback Provision ........ 148
Research Question 4: Beliefs and Practices ....................... 148
Research Question 5: Reasoning Behind Feedback Choices .... 149

CHAPTER FOUR: Results .......................................................... 151
Introduction ........................................................................... 151
Research Question 1: Instructor Beliefs ........................................ 152

Amount of Feedback ......................................................... 154
Importance of Feedback ...................................................... 154
Use of Implicit Feedback ..................................................... 155
Use of Explicit Feedback ..................................................... 156
Treatment of Lexical Errors .................................................. 157
Treatment of Morphosyntactic Errors ................................... 157
Treatment of Phonological Errors .......................................... 158
Summary ............................................................................ 159

Research Question 2: Instructor In-Class Feedback Provision ........ 159

Amount of Feedback ......................................................... 159
Feedback Type ...................................................................... 161
Summary ............................................................................ 161

Research Question 3: Consistency of Feedback Provision .......... 161

Lexical Errors ..................................................................... 162
Morphosyntactic Errors ....................................................... 162
Phonological Errors ............................................................ 163
Incorrect Answers ............................................................... 163
Unsolicited L1 Use ............................................................. 164
Summary ............................................................................ 164

Research Question 4: Beliefs and Practices .............................. 165

Comparing NS and NNS Instructors ....................................... 165
Amount of Feedback ..............................................202
Importance of Feedback ........................................202
Use of Implicit and Explicit Feedback ......................204
Treatment of Specific Errors: Lexical, Morphosyntactic,
and Phonological ..................................................206
Summary ..............................................................208

Research Question 2: Feedback Provision ..................209
Amount of Feedback ..............................................209
  Instructor NS/NNS Status ........................................209
  Instructor SLA Education .......................................210
  Instructor Teaching Experience ..............................211
Type of Feedback ..................................................212
  Instructor NS/NNS Status ........................................212
  Instructor SLA Education .......................................213
  Instructor Teaching Experience ..............................214
Summary ..............................................................215

Research Question 3: Consistency of Feedback Provision ...215
Lexical Errors .......................................................215
Morphosyntactic Errors .........................................216
Phonological Errors ................................................217
Incorrect Answers ...............................................218
Unsolicited L1 Use ...............................................218
Future Research with Current Data ………..…………..243
Future Research with New Data …………………….......….…..244
Summary …………………………………………………..245

REFERENCES ……………………………………………………….246
LIST OF APPENDICES

Appendix A. Participant Recruitment Email ........................................275
Appendix B. Participant Consent Form ..............................................277
Appendix C. Participant Background Questionnaire ...............................279
Appendix D. Stimulated Recall Protocol Questions .................................286
Appendix E. Participant Belief Questionnaire .......................................287
LIST OF FIGURES

Figure 1. Study Design .................................................................101
Figure 2. Amount of Feedback ......................................................160
Figure 3. Model of Traditional Interactive Feedback Research ..........236
Figure 4. Model of Recent Trends in Interactive Feedback Research ........237
Figure 5. Proposed Model of Interactive Feedback Research ............238
LIST OF TABLES

Table 1. Studies Utilizing NS and NNS Instructors ............................37
Table 2. Studies Comparing More and Less Experienced Lang. Instructors 67
Table 3. Correlations for the Three Experience Operationalizations ..........96
Table 4. Participant Background Information .....................................97
Table 5. Linguistic Foci of Recorded Lessons .................................103
Table 6. Stimulated Recall Protocol Participants ...............................107
Table 7. Belief Questionnaire Topics .............................................115
Table 8. Pilot Studies of the Research Design .................................117
Table 9. Dissertation Pilot Study Results: Overall Error Correction ........124
Table 10. Dissertation Pilot Study Results: Error and Feedback Types ......124
Table 11. Dissertation Pilot Study SR Data: Overall Trends .................129
Table 12. Examples of Coding Procedures for Error Type ....................139
Table 13. Examples of Coding Procedures for Feedback Type ...............141
Table 14. Belief Topic and Corresponding In-Class Data .....................149
Table 15. Participants Organized by All Three Independent Variables .....151
Table 16. Belief Topic Means and Standard Deviations .......................153
Table 17. Belief Questionnaire Topic 1: Amount of Feedback ...............154
Table 18. Belief Questionnaire Topic 2: Importance of Feedback ...........155
Table 20. Belief Questionnaire Topic 4: Explicit Feedback ....................156
Table 21. Belief Questionnaire Topic 5: Treatment of Lexical Errors ........157
Table 22. Belief Questionnaire Topic 6: Treatment of Morphosyn. Errors ..... 158
Table 23. Belief Questionnaire Topic 7: Treatment of Phonological Errors 158
Table 24. Amount of Feedback and Instructor ID Variables ....................... 160
Table 25. Significant Results in Chi-Square Tests of Feedback Consistency 165
Table 26. NS and NNS Instructor Beliefs and Practice ......................... 167
Table 27. Beliefs and Practice of Instructors With and Without SLA Ed. ..... 170
Table 28. Beliefs and Practice of Instructors According to Experience .......... 173
Table 29. Stimulated Recall Quantitative Data ........................................ 176
Table 30. Factors Cited in NS Instructor Stimulated Recalls .................... 178
Table 31. Factors Cited in NNS Instructor Stimulated Recalls ............... 182
Table 32. Factors Cited in + SLA Instructor Stimulated Recalls ............... 188
Table 33. Factors Cited in - SLA Instructor Stimulated Recalls ............... 191
Table 34. Factors Cited in + Experience Instructor Stimulated Recalls ....... 194
Table 35. Factors Cited in - Experience Instructor Stimulated Recalls .......... 195
CHAPTER ONE

Statement of the Problem

The role of interactional feedback has been a critical area of second language acquisition (SLA) research for the past three decades given its facilitative effects on L2 development supported by empirical research. Second/foreign language (L2) learning most often occurs in a classroom setting where there is less contact with the target language (TL), placing additional importance on the amount and type of L2 input and interaction to which learners are exposed (N. Ellis, 2007; Norris & Ortega, 2000; Philp & Tognini, 2009). It is argued that interaction, specifically the feedback supplied during interaction, provides opportunities for learners to be alerted to an error in their speech when the learner is already engaged in communication (e.g., Long, 1996; Swain, 1995, 2005). Such feedback increases the possibility of the learner noticing the error, making hypotheses about the correct form, and then testing out these hypotheses. For these reasons feedback is a key consideration in many SLA frameworks and strands of L2 research, especially those based on the interaction approach (Gass, 2003; Hatch, 1978; Long, 1996; Mackey & Gass, 2006; Pica, 1994; Swain, 2005).

However, while feedback is generally held to be facilitative of L2 development (e.g., Keck, Iberri-Shea, Tracy-Ventura, Wa-Mbaleka, 2006; Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Norris & Ortega, 2000; Russell & Spada, 2006), and even hypothesized as essential for learning certain difficult structures (Long, 1996, p. 417), there are many variables that influence the extent of its efficacy. Both internal and external factors have been posited to influence the facilitative effect of interactional feedback on L2 development, including individual difference variables of the learners.
who receive feedback, such as proficiency level (e.g., Mackey, 1999; Mackey & Philp, 1998), age (e.g., Havranek, 2002; Mackey, Oliver, & Leeman, 2003), level of developmental readiness \(^1\) (e.g., Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; Philp, 2003), aptitude (e.g., Erlam, 2005), and anxiety (e.g., Sheen, 2006), as well as learners’ noticing and interpretation of feedback (e.g., Mackey, Al-Khalil, Atanassova, Hama, Logan-Terry, & Nakatsukasa, 2007), and working memory (e.g., Erlam, 2005; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Trofimovich, Ammar, & Gatbonton, 2007). External factors that can influence the effectiveness of feedback have also been investigated and include type of feedback provided (i.e., more or less explicit, e.g., Ammar, 2008; Ellis, 2007; Ellis, Loewen & Erlam, 2006; Havranek, 2002; Lyster, 1998; Mackey, et al., 2007; Panova & Lyster, 2002; Sheen, 2006; Varnosfadrani & Basturkmen, 2009), amount of feedback (e.g., Havranek, 2002), feedback target (e.g., Ellis, 2007; Jeon, 2007; Long, 2007; Mackey, et. al., 2007), feedback focus (reviewed in Mackey & Goo, 2007; Russell & Spada, 2006), the number of changes from the learner’s original utterance (Loewen & Philp, 2006), and the mode of feedback provision such as written (e.g., Ashwell, 2000), computerized (e.g., Rosa & Leow, 2004; Sanz & Morgan-Short, 2004), or oral (e.g., Mackey & Philp, 1998). Context has also become a primary focus of feedback research in the past decade. Studies have compared feedback provision in laboratory versus classroom settings (Gass, Mackey, & Ross-Feldman, 2005), compared feedback according to task features (e.g., Révész, 2009, in press; Robinson, 2003, 2005b), and have compared feedback in different instructional contexts (e.g.,

\(^1\) Developmental readiness is the notion that learners must be at the proper stage of development in order to be ready to learn certain structures; if they are not at the necessary stage, feedback and any other attempts at addressing learner errors on the particular structure, will not be successful (for work on developmental readiness, see Pienemann & Johnston, 1987, and colleagues).
Lyster & Mori, 2006; Sheen, 2004). Recently, scholars have called for more classroom-based research exploring feedback in relation to contextual features within the language classroom (Spada & Lightbown, 2009, p. 169), a focus echoed in both cognitive and social perspectives of SLA research (Dörnyei, 2009; Ortega, 2007; Philp & Tognini, 2009; Spada & Lightbown, 2009).

Despite these numerous directions in interaction-based feedback research and the centrality of contextual factors in the social/cognitive debate in the SLA field, minimal attention has been paid to the feedback provider (referred to as the “source” or “interlocutor” of feedback in the literature), and studies in this area have been limited to comparisons of feedback provided to a student (a) by a fellow learner compared with feedback provided by the instructor or researcher (e.g., Mackey et al., 2003); or (b) comparisons of learner-learner dyads with mixed or similar proficiency levels (e.g., Oliver, 2002). Only two studies have compared feedback provided by instructors according to individual difference (ID) factors (Mackey et al., 2004; Polio et al., 2006). The paucity of research investigating the potential influence of the IDs of the interlocutor involved in communication is surprising given the key role played by the interlocutor during interaction, particularly with respect to feedback provision. In his oft-cited (1996) Interaction Hypothesis, Long explicitly states that:

Negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways (pp. 451-452; my emphasis).
The interlocutor, perhaps more than any other individual, controls the interaction and, even more importantly, the feedback provision believed to be critical for L2 development to occur. Research on the interlocutor is even more compulsory when one considers the influence the most common interlocutor, the foreign language (FL) instructor, has on feedback provision to students on a daily basis. In many cases, the FL instructor is the primary if not the only source of L2 input (Philp & Tognini, 2009). And while several studies have compared feedback provided by students and instructors as well as native (NS) and non-teaching nonnative speakers (NNS), there has been very little research investigating individual differences of language instructors, and how these differences could potentially relate to instructors’ feedback provision in the FL classroom, their beliefs about feedback, and how they make their in-class feedback decisions.

In fact, the vast majority of NS and NNS instructor research undertaken thus far has been restricted to investigations of instructors’ perceived differences (e.g., how NNS instructors believe they differ from NS instructors) between NS and NNS instructors in the English as a Second Language (ESL) or English as a Foreign Language (EFL) contexts (Llurda & Huguet, 2003; Reves & Medgyes, 1994; Samimy & Brutt-Griffler, 1999), as well as examinations of perceived differences by ESL and EFL students (how students believe NS and NNS instructors differ) (Barratt & Kontra, 2000; Benke & Medgyes, 2005; Callahan, 2006; Lasagabaster & Sierra, 2005). While these studies report perceived differences between NS and NNS instructors, there is a lack of empirical evidence demonstrating that these perceptions are correct and that NS and NNS instructors do in fact differ in the classroom the ways they are believed to. Very few studies have compared NS and NNS instructor behavior in the classroom, and most have
focused on instructors’ use of the L1 in the classroom (e.g., Macaro, 2001, 2005); only one study directly investigated feedback (Árva & Medgyes, 2000), and it contains significant limitations. Two studies have begun to explore additional instructor individual difference (ID) factors, instructor education and experience, in relation to feedback provision in ESL learning contexts (Mackey et al. 2004; Polio et al., 2006). However, to date there has been no thorough examination of instructor ID factors in relation to naturally occurring feedback in the foreign language (FL) classroom.

There are several critical gaps in SLA research that will be addressed in this dissertation. First, there has been no investigation of potential differences between the beliefs of instructors of differing NS/NNS status, SLA education and teaching experience, and their provision of oral corrective feedback in the FL classroom, where students receive the majority of their target language (TL) exposure. As Philp and Tognini (2009) state, “in spite of [FL classrooms] being the experience of many L2 learners, the particularities of interaction in FL contexts are largely ignored in research” (p. 246). And while universities across the United States and abroad employ both NS and NNS instructors of various education and experience, potential differences in the way instructors provide feedback in the classroom have been minimally examined. As the use of feedback by instructors according to ID factors has been found to differ outside the classroom, investigating instructors’ IDs in relation to feedback use in the classroom is a logical and necessary next step in feedback-based interaction research. In doing so, this dissertation bridges the gap between early interaction research comparing the feedback provision of non-teaching NSs and NNSs and current strands of interaction research.
investigating ID factors in relation to feedback from a viewpoint too often overlooked: from that of the foreign language instructor.

The contents of the dissertation have been organized into five chapters. The remainder of chapter one defines key terms employed in subsequent chapters. The second chapter begins with a presentation of the theoretical underpinnings of the current study, followed by overviews and reports of empirical studies addressing the variables investigated within interaction-based feedback research to date. The literature review of each variable is followed by a critique of the studies that have explored the variables of instructor NS/NNS status, education and experience. Research questions and hypotheses complete this chapter. The third consists of a comprehensive description of the research design and methodology employed to address the research questions and hypotheses, including a description of the operationalizations, participants, materials, procedures, coding, and statistical procedures. The fourth chapter presents the quantitative and qualitative results from the study in relation to the five research questions. Finally, the dissertation concludes with a discussion of the results in relation to the research questions and hypotheses within the field of interaction-based SLA research, followed by a section outlining the implications for the current study on SLA theory, methods and teacher education, as well as limitations and areas for future research.

Definition of Terms

**Beliefs**: “Convictions of opinions that are formed either by experience or by the intervention of ideas through the learning process” (Ford, 1994, p. 315)… “expressed as
evaluations of what 'should be done,' 'should be the case,' and 'is preferable.” (Basturkmen, Loewen, & Ellis, 2004, p. 244).

**Corrective Feedback/Error Correction/Negative Feedback**: Error correction refers to an error being addressed with negative, reactive feedback (defined below). In the present study, it is synonymous with negative or corrective feedback.

**Explicit Feedback**: Feedback that provides the learner with the correct form while simultaneously indicating that an error was committed (Ellis, 2008).

**Implicit Feedback**: Feedback that alerts a learner indirectly that something in their speech was ungrammatical or otherwise non-target-like (Ortega, 2009).

**Individual Differences/Factors**: Individual differences (ID) refer to “dimensions of personal characteristics that are assumed to apply to everybody and on which people differ by degree” (Dörnyei, 2006).

**Interlocutor**: A person involved in a dialogue, conversation, or oral task.

**Native Speaker**: In this study, a native speaker (NS) refers to someone who learned the language being taught as a primary, or one of several primary languages. A nonnative speaker (NNS) did not learn the language being taught as a primary language.
Oral Corrective Feedback: In this dissertation, oral corrective feedback (OCF) is operationalized as “the reactive information that learners receive regarding the linguistic and communicative … failure of their utterances,” (Mackey, 2007, p. 14). It can refer to implicit or explicit negative feedback, but must follow a learner’s ungrammatical or otherwise non-target utterance.

Practice: The term “practice” refers to behavior in the current study.

Stimulated Recall Protocol: Stimulated Recall (SR) Protocols is a retrospective introspective methodology aimed at uncovering what participants were thinking at an earlier time. Individuals are presented with a stimulus (most often a video-recording of the earlier interaction) and asked to recall what they remember thinking at the time of the original event (Gass & Mackey, 2000).
CHAPTER TWO
Theoretical Background

Without doubt, interactional feedback is viewed by many researchers as central to L2 learning particularly in light of the limited amount of time, input and opportunities for interaction learners have with their target language (TL), unlike when learning their first language (L1) (N. Ellis, 2007; Norris & Ortega, 2000; Philp & Tognini, 2009). The potential benefits of interactional feedback are invoked in numerous SLA theories including the interaction approach, which subsumes the Interaction Hypothesis (Long, 1996), aspects of the Input Hypothesis (Krashen, 1985), the Output Hypothesis (Swain, 2005) and, less directly, Schmidt’s (1990, 2001) noticing hypothesis (Gass & Mackey, 2007). As the current study examines natural interaction in communicative foreign language classrooms, the interaction approach, with its primary focus on meaning-based communication, will be the sole approach discussed in this dissertation.

The Interaction Hypothesis

The nucleus of the interaction approach is Long’s (1996, 2006) Interaction Hypothesis, which maintains that both input and interaction are necessary for language learning to take place, particularly when learning an L2; while input is necessary, it is not sufficient on its own for language learning to occur, as previously argued by Krashen (1985). Interaction, particularly the feedback within interaction, is believed to be a necessary component of language learning for several reasons: First, it alerts learners of errors in their speech during meaning-based interaction. Second, it assists learners in noticing mismatches between their interlanguage (IL) and the target language (TL). Third, it encourages learners to hypothesize the correct forms and test those hypotheses.
and, finally, interaction leads learners to modify their IL and their output, a process considered necessary for language development to take place (Long, 2006).

The Output Hypothesis

The modification of output during interaction is the central tenet of the Output Hypothesis, which claims that learners need to be pushed to use their linguistic resources, reflect on their output and consider ways to modify it in order to learn an L2 (Swain, 2005). Output is believed to assist in language development as it provides the learner with the opportunity for meaningful practice of their linguistic resources, assists learners "noticing the gap" between their IL and their TL, and offers them the chance to test their IL hypotheses (Swain, 2005). Learner output production often brings the opportunity to receive reactive (and often negative) feedback from interlocutors, supplying learners with key information about the linguistic and communicative success or failure of their production (Gass & Mackey, 2007). Some argue that the fact that interactional feedback is provided at a moment when learners are already engaged in meaningful communication increases the likelihood that they will notice the error as their attention is focused on the interaction (Long, 2006). Noticing is a critical component of all aspects of the interaction approach, and is articulated explicitly in the noticing hypothesis.

Schmidt’s Noticing Hypothesis

Schmidt’s noticing hypothesis maintains that noticing input is crucial for language learning to take place, and argues that a learner must be aware of input in order to internalize it (Schmidt, 2001). Noticing is posited to be one of the critical components to the learning process, and researchers have described is as a mechanism for mediating between input and learning (Gass & Mackey, 2007). Simply put, if learners do not notice
their errors, the gap between their IL and the TL, and the correct form, they will not progress in learning the L2. Interaction-based feedback is believed to assist learners to notice their problematic utterances and to allow them the opportunity for L2 development to occur via hypothesis testing and output modification.

Summary

While feedback is theoretically posited to be extremely beneficial for L2 learning for the reasons outlined above, the extent of its effectiveness depends on the influence of various factors, both internal and external. And while the language instructor is inherently important in the approach due to the fact that the language instructor is the most common interlocutor (i.e., the one responsible for interaction-based feedback provision), to date there has been minimal focus on the instructor interlocutor from a theoretical perspective. In the Interaction Hypothesis, Long highlights the importance of interactional feedback stating that,

Negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the NS or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways (pp. 451-52; my emphasis).

This statement has been used as the theoretical justification for interaction-based feedback research for the past two decades. The current dissertation continues and extends the work begun by studies such as Mackey et al., 2004 and Polio et al., 2006, focusing not only on the interactional adjustments (i.e., the feedback) present in interaction, but the relationship between the feedback provided and internal factors of the most common interlocutor: the foreign language instructor.
The next section provides an overview of the variables investigated within interaction-based feedback research to date, followed by a critique of the studies that have explored the variables of instructor NS/NNS status, education and experience.

CORRECTIVE INTERACTION-BASED FEEDBACK

Overview

Numerous factors have been posited to influence the facilitative effect of interactional feedback on L2 development. Many studies have focused on internal factors, such as individual difference (ID) variables of learners who receive feedback, including proficiency level and developmental readiness (e.g., Ammar, 2008; Ammar & Spada, 2006; Iwashita, 2003; Lin & Hedgcock, 1996; Long, Inagaki, & Ortega, 1998; Mackey, 1999; Mackey & Philp, 1998; Philp, 2003; Oliver, 2000; Oliver, 2002), learner age (e.g., Havranek, 2002; Mackey & Oliver, 2002; Mackey, Oliver, & Leeman, 2003; Oliver, 2000), anxiety (e.g., DeKeyser, 1993; Révész, in press; Sheen, 2008), learners’ motivation and attitudes towards feedback (e.g., DeKeyser, 1993; Kim & Mathes, 2001; Nagata, 1993; Sheen, 2007), and measures of aptitude, most commonly operationalized as working memory (e.g., Mackey, Adams, Stafford, & Winke, 2010; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Ranta, 2002; Robinson, 1997; Sagarra, 2007; Trofimovich, Ammar, & Gatbonton, 2007).

External factors investigated to date include: type of feedback provided (i.e., implicit or explicit; e.g., Ammar, 2008; Ellis, 2007; Ellis, Loewen, & Erlam, 2006; Havranek, 2002; Lochtman, 2002; Lyster, 1998; Mackey et al., 2007; Panova & Lyster, 2002; Sheen, 2006; Varnosfadrani & Basturkmen, 2009), amount of feedback (e.g.,
Basturkmen, Loewen, & Ellis, 2004; Havranek, 2002; Lochtman, 2002; Lyster & Ranta, 1997; Oliver, 1996; Panova & Lyster, 2002), length of intervention (reviewed in Li, 2010 and Lyster & Saito, 2010), feedback target (e.g., Ellis, 2007; Jeon, 2007; Long, 2007; Mackey, Gass, & McDonough, 2000; Mackey et al., 2007), and feedback focus (general vs. specific) (reviewed in Mackey & Goo, 2007 and Russell & Spada, 2006). The context of feedback provision has emerged as a primary focus of interactional feedback research in the past decade. Researchers have compared the mode of feedback provision (e.g., Lai & Zhao, 2006; Sheen, 2010), examined feedback according to task features (e.g., Révész, 2009; Révész, in press; Révész & Han, 2006; Robinson, 2003, 2005; Robinson & Gilabert, 2007), and have compared feedback in different instructional contexts (e.g., Lyster & Mori, 2006; Sheen, 2004). Finally, few studies have examined the feedback provider (often referred to as the “source” or “interlocutor”), although these have been limited to comparisons of feedback present in (a) adult NS-NNS and NNS-NNS dyads (e.g., Adams, 2007; Pica & Doughty, 1985; Porter, 1986; Sato & Lyster, 2007; Van den Branden, 1997; Varonis & Gass 1985; Yule & MacDonald, 1990), or (b) child-adult NS-NNS and NNS-NNS dyads (e.g., Mackey & Oliver, 2002; Mackey et al., 2003; Oliver, 2002), with minimal consideration given to the instructor as interlocutor (Oliver, 2000).

To gain an understanding of the interactional feedback research undertaken thus far, general findings of the internal and external variables cited above will be outlined to synthesize findings and demonstrate gaps in the interactional feedback literature. Before entering into this review, it is important to state that although there has been extensive research investigating the presence and efficacy\(^2\) of feedback in both naturally occurring

\(^2\) The term efficacy when discussing feedback often refers to measures of uptake, which is the learner’s response to feedback. Uptake can refer to a learner’s repair of the error or a lack of repair (Ellis, 2008). For
and experimental contexts, conclusions still have not been reached and thus what is presented is the most recent consensus. Many more studies are needed to isolate and thoroughly examine these factors to truly understand the nature of feedback in SLA. It is also necessary to stipulate that the list of studies above is meant to be illustrative of the realm of feedback research, rather than exhaustive, as is appropriate for the current study. For detailed descriptions on these studies and more, see Ellis (2008) and Mackey (2007).

Internal Factors

Many internal factors (commonly referred to as ‘individual differences’) have been investigated as factors potentially mediating the facilitative effects of corrective feedback (and L2 development in general, which is outside of the scope of the current study). Brief summaries are presented below for individual difference (ID) factors that have been investigated to date. For detailed information the reader is directed to Dörnyei (2006, 2009) and Robinson (2000, 2002). It is particularly important to note that the explorations into the mediating affect of the IDs listed have thus far focused only on the learner; the few studies investigating the interlocutor IDs will be examined in the External Factor section below.

Learner Proficiency Level and Developmental Readiness

Learners’ proficiency level has been investigated as a potential moderating factor in studies testing the efficacy of interactional feedback, although there have not been enough of these studies to warrant inclusion in the meta-analyses present in the field (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaleka, 2006; Li, 2010; Lyster & Saito, discussions on the issues of equating uptake with learning, see studies such as Mackey & Philp (1998), Ohta (2000) and, most recently, Yoshida (2010).
2010; Mackey & Goo, 2007; Russell & Spada, 2006; Spada & Tomita, 2010). Thus far, evidence has shown that learner proficiency level or otherwise measured ability (see discussions on developmental readiness and aptitude below) can be an influence (Ammar, 2008; Ammar & Spada, 2006; Iwashita, 2003; Mackey & Philp, 1998). Some findings show that more advanced learners benefit from implicit feedback (recasts) more than their less-advanced classmates (Ammar, 2008; Ammar & Spada 2006; Mackey & Philp, 1998), whereas students of lower proficiency benefit more from more explicit types of feedback (such as prompts). Findings contrary to this general trend include Iwashita (2003), who found recasts to be equally beneficial for all students, regardless of pre-test performance. Learner proficiency level is further discussed under the section Source of Feedback Provision as it has also been investigated as a variable for feedback provision.

Mackey and colleagues posit that learners must be developmentally ready in order to benefit from corrective feedback (e.g., Mackey & Philp, 1998; Mackey, 1999) and explain differing benefits of recasts based on this concept, stating that learners who were developmentally ready to learn the target form were able to benefit from feedback significantly more than learners who were not.

Learner Age

Meta-analyses by Li (2010) and Lyster and Saito (2010) specifically investigated learner age as a potential factor mediating the effectiveness of feedback. Li, investigating both classroom and laboratory-based corrective feedback studies, found that there was no significant relationship between learner age and feedback efficacy. Lyster and Saito,

---

3 The reader is reminded that only Li (2010), Lyster & Saito (2010), and Russell & Spada (2006) focused exclusively on corrective feedback. Feedback was considered in the others though the foci were on task-based or conversational interaction (Keck et al., 2006; Mackey & Goo, 2007, respectively) or comparisons of instruction (Norris & Ortega, 2000; Spada & Tomita, 2010).
focusing solely on classroom-based studies, found that younger learners benefited more from oral corrective feedback than older learners. In an individual study, Oliver (2000) compared the presence and use of negotiation and feedback in ten child (6-12 years old) and ten adult ESL classrooms and found that teachers were significantly more likely to provide oral corrective feedback to adults, though there were no significant differences in the use of feedback (defined as uptake in the study) for the two age groups. Additional studies examining child learners (e.g., Mackey & Oliver, 2002) have found children to benefit more immediately from feedback compared to results from other studies examining adult learners (e.g., Mackey, 1999). Other studies (e.g., Mackey et al., 2003) have considered age while comparing interaction and feedback in learner-learner and learner-NS dyads, and/or in comparisons of child-child dyads versus pairs of adults (Oliver, 2000); this will also be discussed under Source of Feedback below.

The lack of research investigating learner age in relation to feedback is surprising given the prominent focus that both feedback research and age have had in SLA. Many believe learner age plays a major role in SLA; some have even stated that after a certain age (referred to as ‘age effects’ or the ‘Critical Period’ in the literature) achieving native-like fluency in an L2 is substantially more difficult if not impossible after puberty (see discussion in Sanz, 2005). More studies are needed to investigate age as an independent variable to claim any relationship between learner age and feedback provision or efficacy.

Learner Anxiety

While anxiety has been one of the most studied learner ID factors in terms of overall L2 development (Ellis, 2008), like age, it has been scarcely investigated in
relation to error correction. Both DeKeyser (1993) and Sheen (2008) found that learners with low anxiety benefited from certain types of error correction more than others. In DeKeyser’s study, low anxiety learners benefited more frequent, systematic error correction; Sheen found that low anxiety learners benefited more from recasts (an implicit type of feedback) when compared to high anxiety learners. Révész (in press) found no correlation between presence of interactional feedback and learner anxiety, meaning that regardless of learner anxiety level, there were no statistical differences regarding the amount of feedback in task-based learner interactions for those with high or low levels of anxiety. With these limited and mixed results, many more studies are needed to isolate and investigate anxiety in relation to feedback.

Learner Motivation and Attitudes Towards Feedback

Motivation and learner attitudes have also been investigated as potential mediating factors for corrective feedback, albeit minimally (e.g., DeKeyser, 1993; Sheen, 2007). DeKeyser found that frequent and systematic error correction benefited learners who had low extrinsic motivation, while learners with high extrinsic motivation benefited from infrequent error correction. Sheen (2007) found that learners with positive attitudes towards error correction benefited more from metalinguistic feedback than recasts. Like much of the research investigating learner anxiety, the bulk of the motivation research in SLA has focused on general L2 development but not on specific instructional interventions such as feedback provision (see review in Dörnyei & Ushioda, 2009).

Learner Aptitude and Working Memory

Several decades ago, it was common for research to measure aptitude by means of a questionnaire (the Modern Language Aptitude Test, MLAT). Results were contradictory,
with some finding a link between learners’ MLAT performance and their ability to benefit from feedback (Robinson, 1997), and others finding no relationship between aptitude and benefit of feedback (DeKeyser, 1993).

More recently, working memory (WM) has been suggested as being a component of language aptitude (Miyake & Friedman, 1998; Sawyer & Ranta, 2001; Skehan, 2002). WM is defined as “a mental construct that accounts for how the key processes of perception, attention and rehearsal take place.” (Ellis, 2008, p. 983). WM has become an ID factor of great interest for SLA, as it is posited to assist learners process the input they are exposed to in language classrooms (see discussion in Mackey et al., 2010). Some studies have found that learners with high WM demonstrate lasting benefits from feedback, while those with lower WM benefit from feedback immediately and lose the development over time (Mackey et al., 2002; Sagarra, 2007). Erlam (2005) found that having more explicit instruction mediated the differences of learners with high and low WM, although others (e.g., Ranta, 2002) found the opposite result; learners’ WM affected their development regardless of type of instruction. Researchers have also found WM to be significantly related to learners’ use of feedback (i.e., production of modified output, MO): learners with higher WM produced significantly more MO than learners with lower WM (Mackey et al., 2010). Mackey et al. (2002) have suggested that WM may moderate learners’ being able to process verbal input and subsequently be able to benefit from certain types of feedback. Goo’s (2009) study further presses this point; he found that the efficacy of recasts is quite sensitive to WM, but not metalinguistic feedback.
External Factors

As with the research focusing on internal factors, the majority of the research examining the mediating effect of external factors has been limited to considerations of the language learner, with minimal regard for the interlocutor. General trends for each of these factors are discussed below, with more detailed description of the few studies examining the source of feedback provision.

Degree of Explicitness and Feedback Type

As articulated in the theory discussion, noticing is held to be essential in order for L2 learning to take place (Schmidt, 2001). Thus, the degree of explicitness of feedback needed to promote noticing without detracting from the communicative focus of instruction has been and remains a core theme of corrective feedback research (Russell & Spada, 2006). Different types of feedback are most often described theoretically on a continuum of explicitness, rather than as a dichotomy of implicit/explicit (Ellis, 2001, 2008; Mackey & Goo, 2007), though they are most often discussed and empirically tested as dichotomous variables. The most explicit types of feedback are overt corrections, metalinguistic information, and elicitations (Ellis, 2008). On the less explicit end of the feedback continuum (most often referred to as “implicit feedback”) are recasts, confirmation and comprehension checks, clarification requests, elaborations, simplifications, repetitions, and negotiations (Gass & Mackey, 2007). All of these feedback types will be defined in detail and presented with examples later in the paper.

Studies comparing implicit and explicit feedback have fallen into two main categories: (1) studies comparing the efficacy of different types of feedback (i.e., which types promote more L2 development, either measured via pre-, post- and delayed post-
test designs testing a specific linguistic target, or via measures of learner uptake/immediate modified output); and (2) studies describing the natural occurrence of feedback. The former studies have largely taken place in the laboratory setting (particularly those measuring L2 development in a way other than learner uptake) and the latter studies often in the classroom.

Although there have been inconclusive results regarding the efficacy of different types of feedback, meta-analyses demonstrate that, in general, interaction with feedback is more beneficial than interaction without feedback (Keck et. al., 2006). One recent meta-analysis found that explicit feedback tends to be more effective than implicit feedback (Li, 2010), at least immediately and shortly after feedback provision, and others reported inconclusive results comparing explicit and implicit feedback (Lyster & Saito, 2010, p. 283; Mackey & Goo, 2007, p. 440; Russell & Spada, 2006, p.154). The effects of implicit feedback are believed to increase after treatment, at times even bypassing the beneficial effects of explicit feedback (Li, 2010). Researchers speculate that explicit feedback might be more beneficial for more complex features such as difficult contrasts between the L1 and L2, whereas implicit feedback might be effective for less salient features of language learning that are learned and proceduralized over time (Li, 2010; Nicholas, Lightbown, & Spada, 2001). Experimental research regarding feedback type has gradually narrowed in scope and rather than focusing solely on comparisons between implicit and explicit types of feedback, researchers are working to isolate the features within specific feedback types to identify what it is about feedback that is beneficial for L2 learning. This research has been mostly limited to the study of recasts (e.g., Leeman, 2003; Loewen & Philp, 2006; Sheen, 2004); more studies are needed to investigate the
specific features of other feedback types, and to see how IDs and additional factors (e.g., learning context, length of interaction, etc.) can mediate the relationship between feedback and L2 development.

In addition to the numerous studies investigating feedback efficacy, many descriptive studies have taken place in the L2 classroom and have investigated the types of feedback that occur (e.g., Ammar, 2008; Lyster & Mori, 2006; Sheen, 2004). Overall, these studies have found recasts to be the most popular type of feedback used in L2 classrooms (e.g., Lyster & Mori, 2006; Oliver & Mackey, 2003; Seedhouse, 1997; Sheen, 2004; Yoshida, 2008), although recasts have also often been found to be the least effective type of feedback in the classroom when compared with more explicit techniques (Ammar, 2008; Ammar & Spada, 2006; Ellis, 2007; Ellis, Loewen, & Erlam, 2006; Havranek, 2002; Lochtman, 2000; Lyster, 1998, 2004; Mackey et al., 2007; Panova & Lyster, 2002; Sheen, 2006), at least when measured by uptake. However, as the studies took place in a range of instructional (communicative and grammar-focused) as well as linguistic contexts (both second and FL contexts), with minimal attention to additional external and internal factors, more research is needed to isolate and understand how feedback type mediates SLA.

*Amount of Feedback*

Research investigating the type of feedback provision that occurs in language classrooms has reported the overall amount of feedback provision, although it has not been a main focus of feedback research. Lochtman’s (2002) investigation of error correction in German secondary schools revealed that instructors corrected 90% of errors. Lyster and Ranta (1997) investigated feedback in a primary immersion school setting and
found the overall percentage of correction to range from 49% to 69% for four instructors, with an average of 62% of all errors being addressed with oral corrective feedback. In an adult ESL setting, Panova and Lyster (2002) found that instructors addressed student errors with feedback an average of 48% of the time. Oliver (1996) found that instructors provided more feedback to adult ESL learners (addressing an average of 60% of learner errors with corrective feedback) when compared to the amount of feedback they provided to child ESL learners (52%). Basturkmen et al. (2004) explored the error correction of three ESL instructors, finding that more than half of student errors were addressed with feedback (ranging between 51.6 and 82%). Finally, Lyster and Mori (2006) investigated naturally occurring interaction in French and Japanese primary immersion classrooms and found that an average of 67% and 60% of errors, respectively, were addressed with oral corrective feedback. Thus, rates of feedback correction in language classrooms have been found to vary between 48% and 90%. More research is needed to uncover the motivation for such variance, including investigations into internal and external factors of the classroom, and in relation to the learners and instructors involved in communication.

Many of the studies failed to report to overall rate of correction or percentage of correction for specific errors type, instead focusing on the breakdown of the percentage of feedback that was a specific type (e.g., recasts) and the rate of learner repair (Havranek, 2002; Murano, 2000; Sheen, 2004). Others have broken down the amount of feedback by error type (i.e., percentage of feedback that addressed vocabulary, grammar, etc.), but again failed to provide detail regarding the amount of feedback provided overall or for a particular error (Loewen, 2003). In the aforementioned study, incidental focus on form was studied in 12 ESL classes. Results demonstrated that instructors corrected
vocabulary (lexis) the most, 42% of the time, followed by grammar (morphosyntax), 33.3%, and pronunciation (phonology), 21.9%. However, considering that the data collected were collapsed from 12 different classrooms, each with a separate instructor with a range of experience (1-16 years) and qualifications (master’s degree in applied linguistics, CELTA certification, etc.), as well as inconsistent numbers of students in each class (6-12), one is unable to draw conclusions from this one study.

In light of the research examining the type and context of feedback provision discussed throughout this chapter, the lack of research investigating the overall amount of feedback provided in language classes is surprising. If the facilitative affect of feedback is moderated by so many factors, one would assume that the overall amount of feedback provided would seem to be a particularly important variable to study, especially considering the differences reported in the studies above.

Setting of Feedback Provision

While the majority of studies concentrate on either the classroom or laboratory contexts, the presence and effects of feedback have been compared between both settings. In general, corrective feedback in laboratory contexts has been found to be more facilitative of L2 development than corrective feedback in classroom settings (Li, 2010; Mackey & Goo, 2007), although some studies have found little or no differences (Gass, et al., 2005; Russell & Spada, 2006). This is most likely due to several factors: (1) less distractions are present in laboratory contexts when compared to FL classrooms; (2) feedback in laboratories is generally focused on one linguistic target whereas in language classrooms the focus is often on more than one linguistic target per lesson and often on general communication; (3) in-class feedback is generally directed at the entire class,
rather than the one-on-one feedback usually employed in laboratory studies; (4) intermediating factors are better controlled in laboratory studies; (5) measurements of production and learning in laboratories are often more discrete and include grammaticality judgment tests and closed post- and delayed post-tests, whereas measures of production and learning in classrooms are often measured by learner modified output and uptake, both of which have varying definitions and ways of measurements and remain controversial in the field, as previously discussed; and finally, (5) participants in laboratory studies are aware of the fact that they are involved in a research study and thus their attention is likely heightened, increasing the chance they will notice the feedback (Keck et. al., 2006; Mackey & Goo, 2007).

Feedback in FL settings has also been found to be more beneficial than feedback in second language (SL) settings (Li, 2010; Mackey & Goo, 2007). This is believed to be the case for several reasons: studies have shown that learners in FL settings have a more positive view of error correction (Loewen et al., 2009), perhaps playing a role in the efficacy of feedback. Some have found that more feedback (in the form of recasts) is provided and used by learners in EFL settings when compared to ESL classes (Liu, 2007; Sheen, 2004, as cited in Li, 2010). More research is needed to investigate if and how the amount of feedback or student views of error correction relate to feedback efficacy, especially considering that Lyster and Saito (2010) found no significant differences between feedback in FL and second language (SL) classrooms.

Studies have also investigated the presence and efficacy of feedback across FL classroom settings and have found significant differences for presence and use of feedback (i.e., modified output). Both Sheen (2004) and Lyster and Mori (2006)
investigated the presence and use of corrective feedback in different instructional settings, and although recasts were the most utilized type of feedback, learner use of recasts differed according to context. In both cases, recasts were more effective in grammar-focused classes and more explicit types of feedback were more facilitative of modified output in communicative settings. Lyster and Mori explained the results by putting forth a hypothesis that it was the use of feedback opposite of the general setting of the classroom (i.e., in a communicative setting where the focus is on meaning, explicit feedback focusing on grammar would be most effective, and vice versa; referred to as the Counterbalance Hypothesis). All of these studies have led to an increase in the examination of additional contextual variables, discussed below.

Mode of Feedback Provision

Like the previous variable, the mode of feedback provision has most often been considered on its own, with few comparisons of feedback between modes. Li’s (2010) meta-analysis compared feedback provided face-to-face (FTF) with computer-delivered feedback and did not find differences regarding the facilitative benefit of feedback.

Most recently, feedback provision and learners’ use of feedback (uptake) has been compared in FTF and computer-mediated-communication (CMC) and in written and oral modes. In the former category, Lai and Zhao (2006) examined the type of feedback provided for certain errors in both modes as well as students’ use of the feedback in their subsequently produced output. While the differences were not significant, there was more student repair in the CMC mode, where 69% of the feedback was explicit; only 30% was explicit in the FTF mode. Sheen (2010), on the other hand, was the first and only study to compare feedback provision in the oral and written modes. Sheen’s quasi-experimental
study took place in 12 intact intermediate-level ESL classes and found that the degree of explicitness of feedback was related to learner acquisition of English articles, rather than the mode in which the feedback was provided; the more explicit types of feedback were more effective than the less explicit types. With the popularity of feedback studies independently examined in both the FTF and CMC, and in the written and oral modes, this is another area where many more studies are expected in the future.

**Length of Intervention**

Another factor found to mediate the effectiveness of feedback is the length of the treatment or intervention. Li’s (2010) meta-analysis found that feedback within short treatments (lasting 50 minutes or less) was more effective than feedback within longer treatments (60-120 minutes or 120 minutes or longer). However, these differences could also be attributed to setting, as almost all the shorter treatments were in laboratory settings and the longer ones were classroom studies (Li, p. 38). In Lyster & Saito’s (2010) meta-analysis of oral feedback in classroom studies, the opposite finding was reported: longer treatments were significantly more effective than short-to-medium treatments. This differential length of interaction undoubtedly affects additional factors surrounding the efficacy of feedback provision, including the explicitness of the feedback provided. Future studies would do well to compare the same amount of interaction in classroom and laboratory settings and see if conclusions could be made regarding this factor.

---

4 Several researchers have pointed out that the length of interaction can be an issue when it comes to comparing feedback efficacy. Ortega (personal communication, 09/13/09) has argued that a great amount of feedback provided in a lesser amount of time inherently becomes more explicit in nature.
Feedback Target

Research has also investigated if feedback facilitates learner noticing of feedback targeting certain linguistic structures (Ellis, 2007; Jeon, 2007; Long, 2007; Mackey et al., 2000, 2007; Carpenter, Jeon, MacGregor, & Mackey, 2006) as well as facilitating L2 development of certain linguistic structures (Keck et al., 2006; Mackey & Goo, 2007). Studies have investigated if feedback is more facilitative in alerting learners to certain types of errors (such as morphosyntax) compared to others (such as lexis). Several researchers have found evidence that learners’ perceive implicit feedback (recasts) when phonological and lexical errors are targeted more than morphosyntax (Carpenter, Jeon, MacGregor, & Mackey, 2006; Ellis, Basturkmen, & Loewen, 2001; Han, 2008; Mackey, et al., 2000; Sheen, 2006). As is the case for many of the factors investigated thus far surrounding feedback, in the meta-analyses there were conflicting results regarding the linguistic target of feedback. Mackey and Goo (2007) found that feedback targeting lexical items, at least during the immediate post-tests, was more beneficial than feedback targeting morphosyntax; Keck et al. (2006) found no significant results for linguistic target, and this feature was not investigated in Li (2010). However, while some studies have found immediate significant differences for feedback targeting lexis, they also demonstrate that the benefits of feedback targeting morphosyntax (grammar) increase over time (Mackey & Goo, 2007). Thus, as with the delayed benefits of implicit feedback, perhaps learners need to process feedback targeting more complex grammatical features for a longer period of time, rather than benefiting from it immediately as they do from explicit feedback or feedback targeting lexis.
Focus of Feedback

In addition to exploring the target of feedback, studies have also examined whether or not feedback focus (general versus more specific) is more facilitative of L2 development. While it is hypothesized that feedback targeting specific foci would be more facilitative than feedback with a more general focus, results for the most part do not demonstrate this. Russell and Spada’s (2006) meta-analysis found no significant differences between feedback targeting general as compared to specific foci. Mackey and Goo (2007) found focused feedback to have more immediate effects: short-term delayed post-tests revealed a significant advantage for focused feedback compared to unfocused; however the advantage disappeared two weeks later. In both meta-analyses, the authors caution reading into the findings, as the vast majority of studies included provided feedback on specific targets and thus the lack of differences found is not conclusive.

Task-Related Factors

Li (2010) was the only meta-analysis to examine task type in relation to feedback efficacy. While Li did find that feedback given during mechanical drills was more effective than feedback provided during communicative activities, these differences were not significant. Individual studies investigating task features associated with Robinson’s Cognition Hypothesis (2001, 2005, 2007, 2010) have found that task features can influence the provision and efficacy of feedback (Révész & Han, 2006; Révész, 2009). Révész & Han (2006) investigated task familiarity and task type in relation to recast efficacy and found significant results for both task variables: learners who received recasts during familiar tasks outperformed on oral assessments compared with those who received recasts during unfamiliar tasks; task type in combination with recast provision
was also related to students’ performance. Révész (2009), while addressing some of the methodological limitations of earlier studies, explored the task feature +/- contextual support and found that recasts with contextual support (i.e., the more cognitively complex version) were less beneficial than recasts without contextual support for the development of morphosyntax. Finally, Révész (in press) investigated how task complexity affects learners’ focus on form-meaning connections during task-based work in a classroom setting, and whether the relationship was modulated by differences in three learner affective factors – linguistic self-confidence, anxiety, and self-perceived communicative competence – finding that learners reported paying more attention to form-meaning connections during complex tasks, where there were also more instances of discussing metalinguistic features. As there are limited studies investigating tasks and feedback, Révész stressed the need to continue to investigate multiple task types, different features in task design, and interactive and instructional conditions in the future.

*Feedback Provision According to Dyad*

In the field of SLA there has been some investigation into the hypothesis that students may respond differently to corrective feedback depending on who provides it, though this research has been relatively limited considering the prominence of the interlocutor in the interaction approach in general, and with regard to feedback in particular (see *Theoretical Background*). Some speculate that L2 learners might perceive instructor-provided corrective feedback as more valuable because they are native speakers (NS) or near-native speakers of the target language (TL) and thus possess more knowledge about the TL than fellow students, and a few empirical studies have yielded supporting evidence (e.g., Cathcart & Olsen, 1976; Chenoweth, Day, Chun, & Luppescu,
1983; Mackey et al., 2001). On the other hand, there is the hypothesis that students might be less inhibited working with peers and might benefit from peer-generated feedback (Russell & Spada, 2006). However, as with the other areas of feedback research, studies examining differences in feedback provision according to the source have focused on the learner and non-teaching NSs while largely ignoring the language instructor.

Most often, studies considering the feedback provider have compared either the interactional feedback present in (a) adult learner-learner (i.e., NNS-NNS) and adult learner-NS dyads (e.g., Gass & Varonis, 1985; Long, 1983; Pica & Doughty, 1985; Porter, 1986; Sato & Lyster, 2007; Varonis & Gass 1985; Yule & MacDonald, 1990) or (b) have examined feedback in child-adult dyads (e.g., Mackey & Oliver, 2002; Mackey et al., 2003; Oliver, 2000). Results from the former studies have generally found more negotiation (i.e., feedback) present in learner-learner pairs (Gass & Varonis, 1985, 1994; Long & Porter, 1985; Pica & Doughty, 1985; Polio & Gass, 1998; Sato & Lyster, 2007; Varonis & Gass, 1985; Yule & MacDonald, 1990), followed by learner-NS dyads, and the least amount of negotiation occurring during NS-NS pairings (Long, 1980, 1983; Oliver, 1998; Pica & Doughty, 1985, 1988). In the latter set of studies examining feedback provision and use in dyads consisting of child ESL learners and adult NSs, Mackey and Oliver (2002) found that interaction with feedback assisted learners in developing the target structure (question formation) more than the control group. Interestingly, the researchers also found that the beneficial effects for feedback were more immediate for the child learners when compared with adult learners from an earlier study (Mackey, 1999), demonstrating that the IDs of those involved in communication (in this case, the age of the learners) may affect the benefits of feedback. In an earlier study,
Oliver (2000) also examined interactional feedback in NS-NNS child-adult dyads and found differences in the type and amount of feedback provided. Adult learners provided more recasts (26.9% of the time) than their child NS counterparts (15.4%), who provided more negotiations (25.1%) when compared to adults (20%). As for the amount, there were minimal differences: child learners provided feedback for 40.5% of errors, while adults corrected 46.9% of the time.

Oliver (2002) is the most detailed study to date examining the influence of dyad type exploring the relationship between feedback provision and use with the participants’ NS/NNS status (whether or not they were a NS of English), age (adult compared to child), and gender. Results were similar to trends in the adult dyad studies: negotiation occurred more frequently (and had the most different types of feedback) between the NNS learners than with dyads of mixed proficiency, and the NS-NS pairs produced the least amount of feedback. There were no significant differences in feedback provision according to participant age or gender, though there were interesting anecdotal trends: girls tended to use fewer confirmation checks and clarification requests as they became older while boys tended to use more of them as their age increased. Oliver speculated that perhaps differences in terms of feedback type appear later on when learners are older.

Mackey, Oliver, and Leeman (2003) also compared NS-NNS and NNS-NNS gender-matched dyads of adult and children, but this time also considered the developmental level of the ESL learners; all were classified as being on similar developmental levels for the target structure. In this study, adult NS interlocutors provided significantly more feedback than adult NNS interlocutors; child NS interlocutors also provided more feedback than child NNS interlocutors although this
difference was not significant. With respect to the opportunities for student use of feedback (i.e., modified output), there were significantly more opportunities provided by adult NNS interlocutors as compared to the NSs, however learners did not take advantage of these opportunities more than the opportunities provided by the NS interlocutors. Overall amounts of errors addressed with corrective feedback were lower than in other studies: 47% for adult NS-NNS pairs, 32% for adult NNS-NNS pairs, 42% for child NS-NNS pairs and 39% for child NNS-NNS pairs. Mackey et al. (2003) was seminal in that it demonstrated that interlocutor type can relate to feedback provision and use, and suggested that the value of interaction could depend in part on the interlocutor (p. 59).

Fujii and Mackey (2009) extended this line of research to the EFL classroom, investigating learner-learner interaction during paired tasks. They found very low levels of feedback provision for the two tasks: an average of 7% of errors being corrected in one task and only 13% of errors addressed with oral corrective feedback in the second task, despite the high amounts of utterances with errors present. Feedback types varied according to task, as well as the rates of student use of the feedback provided.

*Feedback Provision According to Recipient*

In addition to the dyad studies mentioned above, Oliver has also pioneered research investigating potential differences of instructor-provided feedback according to IDs of the recipient. In her (1996) study, Oliver compared instructor provision and learner use of corrective feedback in ESL classrooms, finding that instructors addressed more errors committed by adult learners (60%) than child learners (52%), although the differences were not significant. Similarly, adults ESL learners also used the feedback more than the child learners (31% compared to 22%, respectively). Oliver’s (2000) study (in part
explained in the Dyad paragraph above) also found that instructors provided more feedback to adult ESL learners compared to child learners. The researcher also found that instructors provided more feedback to both child and adult learners than the non-teaching NSs in the dyad portion of the study, and that the instructors provided more recasts to child learners than to adult learners, who received more negotiation, the feedback type preferred by NSs for learners regardless of age. However, unlike results from the (1996) study, Oliver (2000) did not find a relationship between learner use of feedback and age (adult versus child).

*Source of Feedback Provision*

As demonstrated by the research above, there is still an overall tendency for the interlocutor in feedback research to be a non-teaching individual, despite the fact that the teacher is the most common interlocutor found in natural FL learning. In the 56 studies utilized in their meta-analysis, Russell and Spada found that the source of feedback was limited to the teachers in 29% of the studies, and no background information was provided about the instructor; the only information reported was that the teacher was not the researcher (who was the source of feedback in 37% of the studies). Only 16% of the studies reported that peer NNS learners provided feedback; unspecified non-teaching NSs provided feedback in the remaining 18% of studies.

With so many studies finding differences of feedback provision and use according to interlocutor, it is necessary for researchers to begin to consider ID variables associated with a common interlocutor, the FL instructor, and see if there are any differences in feedback provision and use potentially affecting L2 development. As Russell and Spada state, "Corrective research has focused on mainly teacher feedback in classroom settings
or on native speaker feedback outside the classroom" (2006, p. 136; my emphasis). Mackey and Goo (2007) also called for more studies investigating the interlocutor, stating that, “It seems logical and has been observed by a number of descriptive studies that the effects of interaction may vary depending on the interlocutor (p. 445).” The researchers cited the interlocutor ID factors as a factor ripe for investigation (following Oliver, 2000 and Mackey et al., 2003), a sentiment echoed by Li (2010).

Li (2010) was the first meta-analysis to consider interlocutor in relation to feedback, comparing studies investigating the efficacy of feedback provided by NSs, language teachers or by the computer. Li found that feedback provided by NSs was significantly more effective than feedback provided by language instructors (presumably nonnative instructors?), and also found no significant difference between computer-delivered feedback and feedback provided by either language teachers or NS interlocutors. However, as one would expect, all NS and computer-based interlocutors were present in laboratory settings, while the “teacher” interlocutor was limited to the classroom context. This confounds the variable with setting and fails to provide a true investigation as to the effect of the interlocutor on feedback provision and efficacy. The lack of reporting in particular leaves the reader wondering if the NSs in the laboratory were language teachers as well, or had any background on language acquisition and feedback provision, if they were simply trained for the task at hand, or asked to interact and provide feedback as they deemed natural. Similarly, no information was provided regarding the researchers who created the computer-delivered feedback; were they language teachers? SLA researchers? Finally, Li offered no additional information regarding the nature of the “teachers” in the classroom studies. This lack of reporting and
consideration of instructor IDs in Li’s meta-analysis is unfortunately representative of standard reporting in the field. While great detail is offered regarding the language learners, research design, materials, coding and statistical decisions involved in SLA studies, little information is provided regarding the interlocutor, whether in the laboratory or classroom setting; most studies report only the interlocutor’s native language. Very few report years of teaching experience, or education, and to date, neither of these additional variables has been thoroughly investigated in relation to feedback provision in the FL classroom. The current study seeks to address these critical gaps.

Summary of Interaction-Based Feedback Research

The review above demonstrates the vast and often inconclusive research investigating factors mediating the presence and effectiveness of oral corrective feedback. Though feedback is believed to play a prominent role in SLA, the body of research has focused on either (a) the language learner: how learner ID factors and external factors affect the feedback provided to learners (and if and how learners utilize it) or (b) comparisons of feedback present in dyads according to the individuals’ native/nonnative speaker status (as well as minimal investigation into learner ID factors such as age, gender, and proficiency level). The lack of studies investigating instructor IDs that have the potential to affect their provision of feedback is a critical gap in SLA research and one that will be addressed in the current study.
INSTRUCTOR NATIVE /NONNATIVE SPEAKER STATUS

Overview

The next section details the research investigating differences between NS and NNS instructors. It includes an overview of the topics investigated and methodologies used thus far, as well as the limitations inherent in the studies.

The majority of studies comparing NS and NNS instructors have been limited to investigations of instructors’ perceived differences between NS and NNS instructors in the ESL/EFL context (Llurda & Huguet, 2003; Reves & Medgyes, 1994; Samimy & Brutt-Griffler, 1999), as well as examinations of perceived differences by ESL/EFL students (Barratt & Kontra, 2000; Benke & Medgyes, 2005; Lasagabaster & Sierra, 2005). All of the studies found that both NS and NNS instructors, as well as NNS students, believe that there are inherent differences between NS and NNS instructors. NS instructors are believed to be authentic, relaxed about grades and error correction, lacking the ability to explain grammar well, talkative and lacking formal teacher training, which was believed to be evident in their lesson planning. NNS instructors, on the other hand, were said to be able to predict which structures would be difficult for students, had extensive training and experience, and were more demanding, thorough and less creative (Barratt & Kontra, 2000; Benke & Medgyes, 2005). However, these studies also found that NS and NNS instructors were not seen as differing in their teaching abilities; rather, some participants saw these differences as inherent characteristics accompanying NS or NNS status. Many participants believed that factors other than instructors’ native language is what influences instructors’ in-class decisions, namely, instructors’ education and teaching experience (Benke & Medgyes, 2005; Cots & Díaz, 2005; Lasagabaster &
Sierra, 2005; Samimy & Brutt-Griffler, 1999). While differences in instructor perceptions and beliefs are important, just if and how these perceptions and beliefs influence actual instructor behavior, namely, feedback provision in the classroom, is particularly critical to SLA research. Only a handful of studies have investigated NS and NNS instructors’ teaching behavior, fewer have investigated feedback provision (Árva & Medgyes, 2000; James, 1977; Porte, 1999; Sheorey, 1986) and only one has investigated the relationship between NS and NNS instructors’ feedback beliefs and in-class provision of oral corrective feedback (Árva & Medgyes, 2000), although it was not the main focus of the study; the others focused on comparing NS and NNS instructors’ correction of isolated errors outside of the classroom in the written mode. The limited number of studies comparing NS and NNS instructors are outlined in Table 1.

Table 1. Studies Utilizing NS and NNS Instructors

<table>
<thead>
<tr>
<th>Study Type</th>
<th>Study</th>
<th>Participants</th>
<th>Procedure</th>
<th>Main Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructors’ self-perceived differences</td>
<td>Reves &amp; Medgyes (1994)</td>
<td>216 NS &amp; NNS English teachers from ten countries</td>
<td>Participants were given a questionnaire asking about their perception of differences between NS and NNS instructors</td>
<td>68% of the instructors (the “majority” of whom were NNSs) perceived differences in the teaching behavior of NS and NNS teachers</td>
</tr>
<tr>
<td></td>
<td>Samimy &amp; Brutt-Griffler (1999)</td>
<td>17 NNS graduate students of TESOL</td>
<td>Participants completed questionnaires, participated in class discussions and in-depth interviews and wrote reflections</td>
<td>90% of participants believed there are differences between NS and NNS instructors, though neither type of instructor was believed to</td>
</tr>
</tbody>
</table>
be superior; differences were attributed to instructor background rather than native language

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Methods</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Llurda &amp; Huguet (2003)</td>
<td>101 NNS primary and secondary school EFL teachers in Catalonia, Spain</td>
<td>Participants were given a questionnaire based on Reves &amp; Medgyes (1994) and discussed the answers with the researchers</td>
<td>The secondary school teachers were much more confident in their abilities as NNS EFL teachers; the primary school instructors held the NS as ideal and reported insecurities</td>
</tr>
<tr>
<td>McNeill (2005)</td>
<td>4 groups of high school EFL teachers in Hong Kong (N = 56). Groups were NNS/expert, NS/expert, NNS/novice &amp; NS/novice</td>
<td>The teachers were asked to make predictions about lexical difficulty in a 600-word general science passage intended for high-intermediate students of English whose L1 was Cantonese; instructors were also asked to justify their decisions. Results were compared with students’ perceptions</td>
<td>Regardless of expertise, NNS teachers most accurately predicted which words their students would consider to be the most difficult. And although neither NS group was able to statistically predict the difficulty of the text, the experts outperformed better than the novices, suggesting that experience and training played a role in their decision-making</td>
</tr>
<tr>
<td>Macaro (2001)</td>
<td>6 NS and NNS student teachers of FL French in</td>
<td>Participants discussed the theoretical considerations of L1 use in the classroom</td>
<td>NNS instructors were found to code-switch more than the NSs,</td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Primary teacher experience</td>
<td>Cots &amp; Díaz (2005)</td>
<td>2 NNS EFL instructors and 2 NS EFL instructors in Catalonia, Spain</td>
<td>NNSs code-switched more and demonstrated a higher level of solidarity with the students through their use of the pronoun “we”</td>
</tr>
<tr>
<td>Instructor error evaluation</td>
<td>James (1977)</td>
<td>20 NNS &amp; 20 NS EFL teachers</td>
<td>NNS teachers marked errors more severely than the NSs, though there was large variation within each group</td>
</tr>
<tr>
<td></td>
<td>Sheorey (1986)</td>
<td>64 NS teachers &amp; 34 NNS teachers of ESL in India</td>
<td>Overall NNS teachers judged errors more seriously than NSs. NSs were more strict with lexis while NNSs judged syntax and morphology more severely than NSs</td>
</tr>
<tr>
<td></td>
<td>Porte (1999)</td>
<td>14 NS &amp; 16 NNS instructors of university-level EFL</td>
<td>NNS instructors were found to be less tolerant of errors than NSs</td>
</tr>
<tr>
<td>In-class feedback provision</td>
<td>Arva &amp; Medgyes (2000)</td>
<td>5 NS EFL instructors and 5 NNS</td>
<td>The majority of beliefs and practices</td>
</tr>
</tbody>
</table>
high school EFL instructors in Hungary compared observed behavior with trends from belief questionnaires administered a few years earlier to a separate sample; NS instructors more tolerant of errors

Studies Comparing Self-Perceptions of NS and NNS Instructors

The first published works on NS/NNS instructor comparison were more opinion pieces than empirical studies and thus created controversy and heated debates in the fields of ESL and EFL; the reader is directed to Medgyes (1992) for the most famous example. Medgyes (1994) claimed that the majority of differences of NS and NNS instructors could be attributed to the discrepancy of level of language proficiency, stating that their behavior would be inherently different; Medgyes was also quick to qualify his statement, stating that both types of teachers can be equally effective.

Reves and Medgyes (1994) tested this hypothesis by surveying 216 NS and NNS English teachers from ten countries (Brazil, former Czechoslovakia, Hungary, Israel, Mexico, Nigeria, Russia, Sweden, Yugoslavia, and Zimbabwe). The researchers posited that not only would NS and NNS English teachers differ in terms of their teaching practice, their self-perception and teaching attitudes, they hypothesized that all of these differences would be due to the instructors’ differing levels of language proficiency (1994, p. 354). The questionnaire, administered via mail, consisted of 23 items; 18 were asked of both NS and NNS instructors and 5 were addressed only to NNSs. The majority of the questions were closed and sought to elicit personal information as well as information regarding common teaching contexts; open-ended questions were meant to
elicit the subjects’ self-perceptions and their opinions in relation to the three hypotheses. Results revealed that 68% of the instructors (the “majority” of whom were NNSs) perceived differences in the teaching behavior of NS and NNS teachers. While 84% of the NNSs said their ability in the target language negatively affected their teaching, the rest, 26%, said it did not. The researchers used this self-reported, indirect way of asking about proficiency level to support their three hypotheses. While a study comparing NS and NNS instructors’ attitudes and teaching behavior is unarguably important in the field of L2 learning, this did not occur in the current study. First of all, since there was an overwhelming majority of NNSs in the sample, a fair comparison is not possible. To indirectly ask instructors if their (lack of) native language ability affects their teaching in a negative way is a question biased against the NNSs. To claim that classroom behavior differs for NS and NNS instructors based on the self-reported beliefs of mainly NNS instructors without empirical or observational support provides limited contribution to NS/NNS research. The largest flaw in the study is the claim that these (perceived) differences in attitude and (perceived) classroom behavior are due to proficiency level; no statistical analyses tested this posited directional relationship.

Nonetheless, the Reves and Medgyes (1994) study was replicated in 1999 by Samimy and Brutt-Griffler, who surveyed 17 NNS graduate students of TESOL at an American university. The goals of the study were threefold: to determine how the graduate students perceived themselves as professionals; if they believed there were teaching differences for NNSs and NS instructors; and if they ever felt their NNS status limited them. The researchers expanded the data collection to include questionnaires, classroom discussions, in-depth interviews with the instructors, and analyses of
autobiographical writings of the participants. Results revealed that almost 90% of the 17 participants did believe there was a difference between NS and NNS instructors’ teaching of English, though participants also believed the differences did not make one group superior in their teaching abilities. Samimy and Brutt-Griffler attributed the variances to cross-cultural differences, positing that the sociocultural factors embedded in Western and Asian societies could have influenced the behavior of the instructors more than their proficiency levels; this claim that instructor IDs other than (or in addition to) native language influence their classroom behavior is increasingly supported by empirical research (Benke & Medgyes, 2005; Cots & Díaz, 2005; Lasagabaster & Sierra, 2005).

Llurda and Huguet’s (2003) study also utilized questionnaires to investigate instructor self-perceptions focusing solely on NNSs, like Samimy & Brutt-Griffler (1999). The questionnaire was administered orally in on-on-one interviews with 101 EFL teachers in primary and secondary schools in Catalonia, Spain. The researchers sought to gain insight into how the subjects perceive their own language skills, how these skills affect their teaching, and how the teaching skills have evolved over time. They also asked about the participant’s teaching ideology, their goals as language teachers, and their position in the NS-NNS debate. Results demonstrated that, overall, secondary instructors reported more confidence in their teaching ability and proficiency than primary teachers. Primary teachers reported more awareness of their improvement over time, and also placed more importance on communicative functions in the classroom, whereas secondary instructors focused more on language structure. Half of the primary instructors reported that they would hire more NSs than NNSs, whereas only a third of the secondary instructors reported a NS hiring preference.
Overall, the studies investigating NS and NNS instructors’ self-perceptions speak to the growing belief that instructor IDs influence in-class behavior important for SLA processes, (e.g., feedback provision). Instructors seem to think so. What about learners?

**Student Perceptions of Instructor NS/NNS Status**

Since 2000, studies, mostly unpublished at the masters or dissertation level, have begun to investigate learner perceptions of instructors’ NS status. Several published studies have also investigated learner perceptions of instructors’ NS status. Barratt and Kontra (2000) reported results from three studies: two in Hungary (involving 116 students and 58 teachers) and one in China (with 100 students and 54 teachers). Most students had both NNS and NS English teachers in the past and were asked to free write about their positive and negative experiences. NS instructors were seen as authentic, were reported to be relaxed about grades and error correction, and were said to lack the ability to explain grammar well. NSs were also said to lack formal training, which was evident in their teaching. NNS instructors, on the other hand, were said to be able to predict which structures would be difficult for students, and were reported to have excellent training and experience. While these specific comments about the instructors seemed to favor NNSs, overall students said both types of instructors had a comparable number of pros and cons. What is particularly relevant for the current study is how students saw education and experience as two of the main differences of NS and NNS instructors. Although these factors are reported in many perception studies, they have yet to be thoroughly studied as independent variables.
Lasagabaster and Sierra’s (2005) study investigated English students’ perceptions about the pros and cons of having a native speaker teacher. 76 university students of various specializations (translation/interpretation, Basque, English, German, and Spanish philology) completed closed questionnaires in class, followed by small group discussions about their questionnaire answers. The questions solicited background information and contained 42 statements about native and non-native language instructors, for which students had to report their agreement with the statements on a five-point Likert scale (strongly agree, agree, neither agree nor disagree, disagree and strongly disagree). After meeting in small groups, learners were asked to summarize the opinions from their group discussion in English. The results were presented as percentages and comments, so unfortunately it was impossible for the claims of the researchers (that students overall would prefer to have both NS and NNS instructors, but if they had to choose one, they would choose an NS) to be validated. Like many studies, students reported positive and negative characteristics for both types of instructors, and the lack of statistical analyses leads to limited contributions to the research field.

Benke and Medgyes (2005) also examined students’ perceived differences in teaching behavior of NS and NNS instructors of English. Questionnaires administered in-class to 422 Hungarian students of English aimed at answering the following questions: (1) In the learners’ judgment, which are the most characteristic features of NS and NNS teachers?; (2) In which aspects of teaching behavior are the differences between the two groups most apparent?; and (3) To what extent do learners’ perceptions correspond to those held by the teachers themselves? All students had been exposed to more than a year of English language instruction by both NS and NNS teachers and were at or above a
lower intermediate level of proficiency. All but one of the \( t \)-test comparisons between perceived teaching behavior of NS and NNS English teachers was significant: learners consistently perceived NNS instructors as more demanding, thorough and traditional (i.e. relying on textbooks). NSs were seen as more outgoing, talkative and casual. Both NS and NNS instructors were found to be equally patient. As in other studies, both instructors were seen as having numerous advantages and several disadvantages each (Lasagabaster & Sierra, 2005; Medgyes, 1994). The researchers claim there is only one question remaining: “To what extent do learners’ perceptions correspond to those held by the teachers?” (Benke & Medgyes, 2005, p. 208). Although that question is interesting and warranted, what is arguably critical to research is to investigate how these perceptions correspond with actual behavior. While it would be enlightening to see how (and if) students and instructors have some overlap in their perceptions, it is even more important for the field of SLA to know if the perceptions are accurate or false: Do NS and NNS instructors differ in their classroom behavior?

**NS and NNS Instructors’ Intuitions in the Target Language**

Studies have begun to investigate NS and NNS instructors’ intuitions in the target language, and how accurate they are in predicting which structures their students will find more or less difficult. McNeill (2005) compared native and non-native instructors’ sensitivity to language difficulty from a learners’ perspective (i.e., Are NS and NNS instructors equally able to predict which lexical items will be difficult for their language learners?). The most interesting and unique part of the study is the way McNeill grouped language teachers into four groups with two variables: NS status and expertise, which he
McNeill hypothesized that teachers’ increased language awareness (i.e., expertise) might encompass the ability to analyze and explain language, as well as the awareness of the relationship between the language and the learner. Thus, McNeill posits that a higher level of language expertise/awareness enables ‘expert’ instructors to understand and take into account the L2 learning process. His ideals are primarily based on Berliner’s (1995) theory of the development of teaching expertise, which will be discussed later in the section Instructor Experience. Four groups of secondary English language teachers in Hong Kong (N = 56) participated in the study; groups were non-native/expert (NNE), native speaker/expert (NSE), non-native/novice (NNN) and native speaker/novice (NSN). The ‘expert’ teachers all had a postgraduate teaching qualification and a master’s degree in applied linguistics/TESOL or equivalent and had been teaching English in Hong Kong for at least two years. Instructors referred to as ‘novice,’ on the other hand, were English subject specialists in the first year of a full-time bachelor’s degree in education at the University of Hong Kong; their only experience teaching English was private tutorials outside of class. While all native speakers of English had some familiarity with Cantonese (the students’ L1), none had studied the language formally.

The teachers were asked to make predictions about lexical difficulty in a 600-word general science passage intended for high-intermediate students of English whose L1 was Cantonese; instructors were also asked to justify their decisions. 200 students read the text and were then given a vocabulary test from words in the passage; their answers were compared with the teachers’ predictions. Results revealed that the two non-native speaker groups, NNE and NNN, outperformed both of their native speaker
counterparts; correlations of both native speaker groups’ predictions were not significant when compared with students’ answers. By contrast, both native speaker groups demonstrated significant correlations with the difficulty level of the words. When the performances of the groups were compared by expertise, both non-native groups (i.e. both novices and experts) produced significant correlations with the student answers. Thus, results suggests that regardless of expertise, teachers who speak their students’ L1 (Cantonese, in this case) had a distinct advantage in knowing which words their students would consider to be the most difficult. And although neither NS group was able to statistically predict the difficulty of the text, the experts performed better than the novices, suggesting that their experience and training played some role in their decision-making. A qualitative analysis of the instructors’ reasoning behind their difficulty choices did not reveal significant patterns according to group but did reveal vast within-group differences. McNeill called for an increase in measures of teaching expertise and additional investigations into the relationship of teaching expertise and language awareness of instructors.

There is one major limitation to the study: McNeill’s operationalization of expertise. While well defined in the study, McNeill failed to justify why he collapsed instructors’ linguistic/language teaching education and years of teaching experience. The researcher stated that he expected both components to influence (increase) instructors’ awareness of the target language and also increase their awareness of “the relationship between the language being taught and particular groups of learners” (p. 110). And while McNeill found differences according to expertise, these differences were not significant; McNeill is unable to claim if it was the combination of education and experience or only
one of the components that was the influence in instructors’ intuitions. The way McNeill used the term ‘expert’ to refer to an instructor with a minimum of two years of experience is also a unique way to operationalize the term; most would consider that person to still be a ‘novice’ teacher (see discussion in *Instructor Experience*). Future studies must take care to separate confounding variables to be able to relate instructor IDs and actual teaching behavior. If McNeill had separated education and experience perhaps he would be able to explain some of the within-group differences, or even find differences reaching the level of significance; due to the confounding nature of his operationalization of expertise, one can only speculate. Future research should also investigate teaching behaviors theorized to greatly affect the SLA processes. The current study addresses these limitations by exploring instructor education and teaching experience in isolation as well as in relation to feedback provision, a classroom behavior believed to be important for SLA.

*NS and NNS Instructors’ Teaching Behavior*

Research examining the potential differences in teaching behavior of NS and NNS instructors is slowly increasing, though it is still greatly limited. Studies comparing NS and NNS teaching behavior have investigated code-switching (Macaro, 2001) and teacher talk (Cots & Díaz, 2005), and have found differences: NNS instructors were found to code-switch more than NSs (Macaro, 2001), and demonstrated a higher level of solidarity with the students through their use of the pronoun “we” (Cots & Díaz, 2005). These few studies have been some of the first to investigate differences of NS and NNS instructors in the classroom. While these studies are a necessary first step, other
important interaction differences, such as feedback provision, need to be addressed. As feedback provision has been found to be variable in naturally occurring speech for non-teaching NSs and NNSs (Gass & Varonis, 1985; Long, 1983; Mackey & Oliver, 2002; Mackey, Oliver & Leeman, 2003; Oliver, 2000; Pica & Doughty, 1985; Porter, 1986; Sato & Lyster, 2007; Varonis & Gass 1985; Yule & MacDonald, 1990, among others), investigating instructor NS/NNS status as a potential influence of feedback provision in the classroom is a logical next step in L2 interaction research, and one study has begun this strand of research (Árva & Medgyes, 2000). Before discussing this study in detail, several studies that have compared NS and NNS error correction outside of the FL classroom (in the written mode) will be discussed to further demonstrate the limitations of previous research and the need for the current study.

**NS and NNS Instructors’ Error Evaluation: Outside of the Classroom**

Several early studies compared NS and NNS instructor error correction outside of the classroom in an area of research known as Error Evaluation (see Ellis, 2008, for a review). James (1977) examined the error correction of 20 NNS and 20 NS EFL instructors by giving participants 50 isolated statements taken from EFL learners’ compositions. Instructors were asked to locate the error in each statement and rate its gravity on a Likert scale. Results revealed that NNS teachers marked errors more severely than NSs, though large variation existed within each group. In Sheorey’s (1986) study, 64 NS and 34 NNS ESL teachers in India were given 20 sentences taken from 97

---

5 There are many additional studies comparing the error evaluation of NNS instructors and non-teaching NSs (e.g., Davies, 1983; Hughes & Lascaratou, 1982; Hyland & Anan, 2006), but as the focus of the current study is on instructor provision of feedback (not error evaluation) in the oral (not written) mode, these studies are omitted.
ESL learners’ compositions and asked to judge the seriousness of the errors. Overall, NNSs judged errors more seriously than NSs. However, NSs were stricter with lexis while NNSs were more severe with syntax and morphology.

Porte (1999) was a more recent study in this area and investigated the mean scores that EFL professors assigned to their students’ academic writing and again found NNS instructors to be less tolerant of errors than NSs. 14 NS and 16 NNS university instructors were asked to respond to student errors taken from a random selection of 54 anonymous un-graded compositions. All instructors had at least five years of EFL experience and at least two years experience grading EFL students' writing. Instructors were given a questionnaire containing 20 sentences, each sentence containing one error, and asked to rate the gravity of the error on a five-point Likert scale. While the study did reveal differences in NS and NNS instructor error toleration, the methodology in this study, like the previous error correction studies, may be questionable at best. Professors were presented with individual sentences isolated and out of context, an entirely different process from day-to-day student compositions; it does not necessarily follow that instructors would react the same way to isolated errors as they would to errors in context. The questionnaire was also potentially problematic as it asked professors to rate errors on a scale of zero to five; zero being no error and five being a serious error. No operationalizations of rankings two through four were given to professors, and no operationalization of ‘serious’ (i.e., completely impedes communication, violates rules of the language, etc.) was explicitly stated, thus limiting the construct validity of the findings. In view of the limited scope of the study, Porte called for more studies.
investigating NS and NNS instructors’ treatment of student errors, particularly in non-ESL/EFL contexts, where the majority of NS/NNS research has occurred.

The methodological limitations of Porte (1999) can be seen throughout error evaluation research. In the vast majority of error evaluation studies, learner errors are usually presented out of context, in isolated sentences taken from learners’ compositions or other work, or created for the purpose of the error evaluation itself. This contrived and unnatural way of presenting learner errors is unlike any pedagogical or real-life context in which error correction would take place and thus the results from the studies mean little outside of the study. Additionally, the majority of EE studies did not ask NSs and NNSs to actually correct the errors, just rate their severity, intelligibility and irritation. Any relationship between these comments and subsequent error correction cannot be claimed.

NS and NNS Instructors’ Feedback Provision Inside the Classroom

To the best of the researcher’s knowledge, only one published study has compared NS and NNS instructor in-class provision of feedback within the foreign language classroom. Árva and Medgyes (2000) examined differences in NS and NNS instructors’ stated beliefs about the classroom habits of NS and NNS EFL instructors with their observed classroom behavior, which included knowledge of grammar, language competence, attitude, formality and humor incorporated into the class atmosphere, textbook use, cultural information and strictness, as well as error correction. The researchers recorded ten hours of classroom lessons by NSs ($n = 5$) and NNSs ($n = 5$) and compared observed behavior with trends from belief questionnaires administered a few years earlier (Medgyes, 1994). Recordings took place in secondary grammar schools.
in Hungary, and observed classes were of a variety of language levels and used a variety of different textbooks. Results, presented anecdotally without statistical analyses or empirical evidence, claimed that the majority of beliefs and practices corresponded: NSs spoke “better” English than NNSs and appeared to be more “flexible” and “innovative;” NNSs were 'stricter' in teaching style but also more likely to 'empathize' with students, having been in their position in the past; NSs were more “tolerant” of errors than NNSs and corrected “less;” NNSs spent more time lesson planning and were more casual in their appearance; NSs used rich cultural examples and designed their own materials; and NNS relied on the textbook and did not use culture to enrich the classroom (p. 364).

Though Árva and Medgyes (2000) attempted to increase research focusing on NS and NNS instructors, the methodological limitations and subjective opinion preclude any substantial contributions to the field of SLA. In terms of methodology, recordings took place in a variety of language levels at different schools that used different textbooks. NNSs often were the instructors of grammar classes, whereas NS instructors taught conversation classes. This confounds the study, as differences in teaching styles could not be solely attributed to NS or NNS status, but also to different course goals. As previously articulated, results presented as comments lacking empirical support were also saturated with subjective judgments that had not been operationalized: “better” English, “stricter” teaching styles, more “tolerant” of errors, etc. The fact that teachers were informed that the study was investigating differences in NS and NNS teaching style prior to the classroom recordings is an additional limitation, as this knowledge could have influenced behavior during the recorded classes and thus fails to provide an actual representation of normal teaching behavior. The researchers also stated that, “our attempt to homogenize
the student sample was only partially successful” (p. 360), without any further explanation of how these classes were manipulated and thus varied from natural classrooms. Finally, the researchers failed to provide any additional background information regarding the training and experience of the language instructors. Thus, the researchers cannot unequivocally state that the observed differences were solely based on the NS/NNS status of the instructors. The study is important, however, in that it was the first of its kind to directly investigate potential differences of NS and NNS instructors in the language classroom and has also been the only study that investigated differences between NS and NNS instructor corrective feedback beliefs and practices. It is also the only belief and behavior feedback study to date in the oral mode. Beliefs are important to investigate as they are one of the factors that can influence instructors’ classroom behavior although, as previously discussed, instructor beliefs are only one of the potential factors worth investigating and should not be equated with actual behavior, as evident by the complex results from teacher cognition research (discussed starting on p. 82).

Strengths and Weaknesses of Existing NS/NNS Instructor Research

Though there are relatively few studies that have investigated instructor native speaker status when one considers the vast field of SLA, there are several positive aspects to the existing research. The majority of studies comparing NS and NNS instructors have focused on instructors' perceived differences and beliefs (i.e., how they believe NS and NNS instructors differ) (e.g., Reves & Medgyes, 1994). Studies comparing NS and NNS instructor beliefs are important to gain insight into potential differences between instructors. As many studies in the SLA literature have demonstrated, instructor beliefs
can influence their classroom behavior (e.g., Tsui, 2003), though it is not always the case that they actually do (e.g., Borg, 2003, 2006a) (discussed in the Beliefs and Practice section below). It is important to investigate these beliefs both on their own, as the majority of studies have done, as well as in relation to observed classroom behavior. The limited number of studies investigating instructor NS status and its relationship with classroom behavior have compared instructors’ use of code-switching (Macaro, 2001), use of teacher talk (Cots & Díaz, 2005; Shin & Kellogg, 2007), and overall teaching style and feedback correction (Árva & Medgyes, 2000). These studies all utilized naturally occurring classroom data (i.e., non-experimental), which has allowed for investigations into natural classroom interaction, providing a valuable contribution to the field of SLA.

There has also been limited research of learner error evaluation in SLA that has compared NS and NNS instructors’ toleration of learner errors (e.g., James, 1977; Porte, 1999; Sheorey, 1986; see also review in Ellis, 2008). These studies were strong empirically in that both NS and NNS instructors graded the same learner errors taken from students’ compositions; the studies also compared NS and NNS instructors’ toleration for specific types of errors (i.e., morphosyntax, lexical, etc.), allowing for detailed analyses and direct comparisons.

While there are evident strengths in the existing instructor NS/NNS status literature in SLA, the limitations far outweigh the positive aspects of the previous research. First of all, the largest limitation is the overwhelming lack of detail in reporting. Interactional feedback research in large part began due to comparison of NS and NNS interlocutor speech in studies such as Gass and Varonis (1985), Long (1983), and Pica (1987, 1988). However, while studies often report details of their learner
participants (i.e., level of proficiency, native language, additional languages studied, etc.),
the majority of studies fail to provide details about the *interlocutors*; usually only
NS/NNS status is reported and even that construct is scarcely defined. This is a critical
limitation, as all NSs and NNSs are not necessarily alike, particularly those with SLA
education and teaching experience and those without, as was seen in McNeill (2005) and
will be discussed further in the following sections. This assumption that all NSs and all
NNSs are alike, and the failure to report additional instructor ID details, is also important
when one considers how many of the studies have observed natural interaction in
language classrooms, or asked the interlocutors to interact naturally in a laboratory
setting, providing feedback whenever they felt it was necessary. Some studies have taken
the assumption that all NSs and NNSs are alike one step further by utilizing belief data
from one population of NS and NNS instructors, and behaviors from another. Árva and
Medgyes (2000) recorded ten hours of classroom lessons by NS and NNS Hungarian
EFL instructors and compared observed behavior with trends from belief questionnaires
administered to a different group of instructors a few years earlier. Studies must also
report and take into account instructor IDs such as instructor education and experience.
Reporting additional instructor ID factors will allow for empirical comparisons between
other NS and NNS instructor populations and will allow researchers to take the additional
factors into account when interpreting results.

The second limitation of the existing research is the lack of contextual variance:
the majority of studies investigating instructor NS status have been limited to the ESL
and EFL environments. English is rapidly becoming a worldwide “lingua franca,” in
addition to an officially recognized language in many nations. With such status,
additional factors such as new dialects and varieties (i.e., “Euro” English) and language programs at national levels continue to separate English from other foreign languages. Thus, more foreign languages must be studied to see if differences according to instructor NS status are present in other L2 contexts. The open comparisons between second and foreign language instructors (again, often of English) also need to be taken into account. There are many differences between second and foreign language contexts, and thus the data is not directly comparable (Li, 2010). An additional contextual limitation of the current instructor NS status research regarding feedback provision is the utilization of de-contextualized, isolated errors (e.g., Porte, 1999). Studies are needed that investigate instructor NS status and feedback provision in the FL classroom, where feedback provision naturally occurs and has the potential to influence the majority of FL learners.

The measurements utilized to obtain results further limit existing research. Many studies comparing NS and NNS instructors actually investigated instructors’ beliefs about how NS and NNS instructors would presumably differ in the classroom (e.g., Reves & Medgyes, 1994); few actually measured behavioral differences. Some of the leading prose in the questions contained in belief questionnaires such as the one used in Reves and Medgyes (1994) further limit the interpretation of the results. Reves and Medgyes’ (1994) questionnaire included questions to NNS instructors that asked if their non-native language difficulties negatively affected their teaching. Questionnaires must take precautions so as not to pose leading or biased questions (Dörnyei, 2010). The Reves and Medgyes (1994) study also demonstrates another limitation of the existing research, which is the often-indirect measurement of proficiency level. Researchers such as Reves and Medgyes often claim that, when NNS instructors answer questions such as the one
above in the affirmative, there is a direct relationship between proficiency level and classroom behavior, which is not necessarily so. Again, data from belief and perception questionnaires must be treated as such, and not used as evidence that differences exist in behavior. Direct measurements of instructor behavior are needed to make such claims.

Another key constraint is the lack of construct operationalization and the presentation of the data to be compared by NS and NNS instructors. The error evaluation studies (James, 1977; Porte, 1999; Sheorey, 1986) investigating instructor NS/NNS status and written feedback gave instructors isolated sentences from students’ compositions and asked them to rate errors on a scale of zero to five. First, the presentation of individual sentences isolated and out of context is entirely different from day-to-day student compositions; it does not automatically follow that instructors would react the same way to isolated errors as they would to errors in context. Having instructors rate errors on a scale of zero to five; zero being no error and five being a serious error, was also problematic, as no operationalizations of rankings two, three, and four were given to professors, and no operationalization of “serious” was provided. Other studies have utilized subjective judgments as evidence of NS and NNS instructor differences, without operationalizing what ‘better’ English, ‘stricter’ teaching styles, more “tolerant” of errors, empirically means (Árva & Medgyes, 2000).

Some studies have also utilized NS and NNS instructors from differing populations, and others have limited their investigation to perceptions and behavior of NNSs while ignoring NSs (e.g., Llurda & Huguet, 2003; Samimy & Brutt-Griffler, 1999). While the aforementioned case studies are valuable in their in-depth analysis of specific instructors, the claims of NS and NNS instructor differences are, again, beliefs of
perceived differences, not actual differences, and to only ask NNSs and not NS instructors but then claim that the research is a comparison of both is not accurate (Samimy & Brutt-Griffler, 1999).

On a similar note, studies have also compared NS and NNS instructors in incomparable contexts and in classrooms with differing linguistic targets. The Árva and Medgyes’ (2000) study, which examined differences in NS and NNS instructors’ stated beliefs with observed classroom behavior, is a prime example. Recordings took place in classes of varying language levels at several schools that utilized different textbooks. Furthermore, the NNSs taught grammar classes, whereas the NSs taught conversation. This confounds the study, as observed teaching differences could not solely be attributed to instructor NS/NNS status, but could be due to other factors or a combination of factors, including course goals, classroom language level, teaching methodology followed, linguistic target of the lesson, etc.

Studies have also violated a cardinal research rule by influencing the participants, sometimes directly, in a way that biases the findings. Some researchers (e.g., Árva & Medgyes) have given belief questionnaires prior to classroom observations, potentially alerting instructors to the research interests before observing classroom behavior. This is a significant problem, as instructors’ classroom behavior could very well have been influenced by their knowledge of the researchers’ interest; thus, the observed classroom behavior might not have been natural. Even more severe are studies that directly informed instructors of the purpose of the study prior to observations, such as Árva and Medgyes (2000); in this study the researchers informed teachers that the study was investigating differences in NS and NNS teaching style prior to the classroom recordings,
which could have influenced behavior during the recorded classes and thus failed to provide an actual representation of normal teaching behavior. Future studies must take care not to influence the observation of instructor behaviors, and must administer belief questionnaires after the observations are completed. Only after all data is collected should researchers inform instructors of the purpose of the study.

In terms of data sample size, the modest numbers of observed instructors are also quite limited; the largest observational study comparing NS and NNS instructors was Árva and Medgyes’ (2000) study utilizing 5 NSs and 5 NNSs. Studies investigating instructor beliefs have had much greater sample sizes; Reves and Medgyes’ (1994) study is one of the largest, surveying 216 English instructors from ten countries. The small sample sizes of the observational studies also limit the claims of the studies’ ability to compare NS and NNS instructors at large, when the investigations reflect the sample size and methodology of case studies. Investigations are needed that examine much larger sample sizes of NS and NNS instructors, particularly studies exploring potential differences in behavior, such as feedback provision, as well as studies comparing beliefs and behavior.

Researchers must also investigate the reasoning behind these NS and NNS instructor differences; belief questionnaires and classroom observations can only do so much. The first step in this line of research is to explore why the differences exist and how they affect the SLA process. In order to truly understand instructor ID factors and how they influence feedback provision, the use of introspective methodology to investigate potential differences is a needed step in interaction-based feedback research.
Finally, studies such as McNeill (2005) have attempted to investigate additional instructor ID factors by comparing NS and NNS instructors according to expertise, as well as native language. However, this was limited to one study and its operationalization of expertise (linguistic education and at least two years of teaching experience) confounds two potential variables. Defining an instructor as an ‘expert’ with only two years of experience is inconsistent with the literature and there was no justification of the operationalization. Future studies must take care to explain and justify each operationalization, and make sure to measure ID variables separately in order to claim any relationship between an ID and behavior. This research is particularly important in light of the fact that so many studies have reported a belief that instructor education and experience influence their in-class behavior (Benke & Medgyes, 2005; Cots & Díaz, 2005; Lasagabaster & Sierra, 2005; McNeill, 2005).

In conclusion, studies investigating instructor NS/NNS status are gaining momentum in the field of SLA, though the many limitations demonstrate the need for more robust research designs. Specifically, future studies need to report instructor ID information in addition to NS/NNS status and ideally take these factors into account when analyzing the data. More studies are needed in the FL classroom setting, with more instructors, and laboratory settings must provide comparison data (e.g., students’ errors from compositions) in a more natural, contextualized way. Future studies must also increase the robustness of the measurements used in instructor NS/NNS status research: belief questionnaires must be utilized solely for measuring beliefs, and in-class behavior observational data for behavioral comparisons. Researchers must also observe first and administer belief questionnaires second, so as not to influence behavior. All constructs
must be operationalized and researchers must remove subjective opinion as coding for comparisons. In order for data to be empirically comparable, NS and NNS instructors must be from the same population, data must be collected from classrooms of comparable goals and teaching methodology, and from lessons of similar linguistic targets. Finally, there is a critical need for more studies investigating instructor NS/NNS status and classroom behavior, particularly in relation to feedback provision. Future research investigating instructor NS/NNS status needs to take all of these limitations into account in order to truly address the question if there are differences attributable to NS/NNS status, or if perhaps the differences found thus far are related to other ID factors, such as instructor SLA education and teaching experience. Studies investigating the former variable, instructor SLA education, are critically discussed in the next section.

INSTRUCTOR EDUCATION

Overview

Research investigating L2 instructor education, like instructor NS/NNS status research, has also increased in the recent years (Burns & Richards, 2009). This research, however, has focused principally on three areas: (a) how instructors’ training courses affect their opinions and knowledge about teaching and/or about the target language (e.g., Almarza, 1996; Bayliss & Vignola, 2007; Borg, 2005; Cabaroglu & Roberts, 2000; da Silva, 2005; Flowerdew, 1998; Freeman, 1993; Gutierrez Almarza, 1996; Kagan, 1992; MacDonald, Badger, & White, 2001; McDonough, 2006; Pennington & Urmston, 1998; Richards, Ho, & Giblin, 1996; Sendan & Roberts, 1998; Urmston, 2003); (b) examinations of the types of teacher knowledge and how they affect teaching practices
(e.g., Almarza, 1996; Andrews, 1997; Andrews, 1999b; Bartels, 1999; Borg, 1998b, 1998c; Gatbonton, 1999, 2008; Golombek, 1998; Spada & Massey, 1992); and (c) what language teachers should know (e.g., Freeman, 2004; Freeman & Johnson, 1998; Johnson, 2009; Yates & Muchisky, 2003). Very minimal research has investigated if instructors’ second/foreign language acquisition education (beyond a teacher training course) relates to their beliefs (McDonough, 2006; Peacock, 2001) and in-class behaviors believed to influence the SLA process. To the best of the researcher’s knowledge, none have investigated the relationship between instructor SLA education and in-class corrective feedback; only one study investigated the influence participation in a workshop had on teacher use of incidental focus on form (Mackey et al., 2004).

**Existing Research on Instructor Education**

Overall, the aforementioned studies have shown that instructor beliefs and knowledge do change as a result of training courses, although it is outside the scope of the current study to discuss this issue in detail. Research investigating pre-service teachers has revealed that, prior to formal teacher education classes, instructors base their theories of language teaching on their personal experiences as language learners (Bailey, 1996; Farrell, 1999; Johnson, 1994; Numrich, 1996; Urmston, 2003; Warford & Reeves, 2003). After teacher education courses, instructors have been found to change their beliefs about language teaching and learning to reflect what they have learned (see reviews in Borg, 2003, 2006a). The vast majority of these studies investigated changes taking place after a one-semester course or certification (such as TESOL or CELTA, Certificate in English Language Teaching to Adults); only a couple investigated changes
after participation in longer courses. Peacock (2001) was one of these studies.

Peacock (2001) measured the beliefs of 146 ESL teachers at the beginning of their training course and again after 2 years and 3 years of study of TESOL methodology. The instrument used in the study was Horwitz’s (1985) Beliefs About Language Learning Inventory (BALLI), which focuses on beliefs about how languages are learned and how they should be taught. In this study, unlike many of the others, beliefs did not change after education. Peacock called for more research investigating instructors’ beliefs over time, especially in the context of longer teacher education coursework, to see which factors in instructor education affects beliefs and which are not facilitative of change.

Although only a semester in duration, McDonough’s (2006) study is important to mention, as it is to date the only study that investigated the influence of a graduate-level action research seminar on language teaching. Seven graduate teaching assistants participated in the elective seminar, where they discussed articles, kept a journal, wrote reflective essays, participated in a focus group and carried out an action research project in their language classrooms. McDonough, using a qualitative inductive research approach, found that instructors’ opinions and self-reported classroom behaviors changed as a result of the seminar, and the majority of these changes were lasting; all reported that the seminar changed their teaching when the researcher contacted them 13 months later.

As mentioned in the introduction, Mackey et al. (2004) was the only study to investigate the effect of teacher education on feedback provision. The follow-up study in the article examined the effect of a workshop on four inexperienced ESL instructors’ use of incidental focus on form (FonF). The researchers videotaped two lessons of each instructor prior to the workshop and compared teachers’ FonF provision in two lessons
after their workshop participation. Findings determined that instructors’ use of incidental FonF did not change, despite having participated in the workshop. However, instructors’ reflections demonstrated an increase in awareness and thought about incidental FonF use in the classroom. Mackey et al. conclude by stating “These studies suggested that investigations into incidental focus on form should take into account individual differences in terms of L2 teachers’ experience and education” (2004, p. 321). The researchers highlight the need to investigate why these differences exist, citing the possibility of using retrospective measures such as stimulated recall, to uncover a deeper understanding into instructors’ use of reactive FonF, particularly in relation to certain linguistic targets in natural classroom settings. Finally, the researchers acknowledge the brevity of the workshop that constituted instructor education and the need for additional research investigating instructor education in relation to feedback provision.

Limitations of Existing Research/Motivations for Investigating Instructor SLA Education

There are many areas to improve upon in the area of instructor education research. First, future research should investigate the influence of instructor education in the area of applied linguistics and SLA. Rather than focus on language teaching in general, which often focuses on classroom planning and practices rather than SLA processes (Ellis, 2009), it would be interesting to see if instructors’ SLA education influences their beliefs and behaviors. Although teaching methodology courses often discuss language learning processes briefly (Burns & Richards, 2009), they are also required to cover myriad of topics such as lesson planning, assessment, teaching methods and approaches, error correction, task and activity creation, differences in teaching content compared to
grammar, etc., as well as the rules and expectations for a certain department or school. SLA courses, on the other hand, are able to focus on language processes and go into depth about learner affective factors, the different theories on how languages are learned, the cognitive processes involved, contextual considerations, what the research literature shows in detail. The researcher would argue that a background in SLA theory and research is inherently different from teaching methodology courses and is equally if not more important to consider as an instructor ID factor.

Second, by investigating instructors’ SLA education, one would eliminate an inherent limitation pointed out by Borg (2006a, p. 65): by assessing instructors’ beliefs and behaviors at the beginning and end of a training course, the participants know they are being assessed for change that occurred presumably as a result of the course. As they are students who will be receiving a grade and, often, going on to teach in the same language department, this could put pressure on the participants to respond in a way they thought was expected, rather than in a way truly reflecting their beliefs and behavior. Examining SLA background, on the other hand, would not bias the data or create any pressure on the participant to answer in a certain way, as the instructors’ coursework is not in any way tied to the data collected.

Third, as differences have been found after a one-semester course, it seems logical to assume that instructors would be even more affected by more substantial coursework (for example, in SLA, as is the focus of the current study). While most teacher training courses last at most one semester, investigating a certain aspect of instructors’ educational background, such as SLA education, would expand the research of teacher education and allow studies to target potential differences in beliefs and behaviors from
instructors who have participated in SLA coursework and those who have not.

Fourth, studies investigating instructor education have been greatly limited to general teaching beliefs and processes and have neglected to investigate potential relationships between instructor SLA education and provision of feedback. Researchers must investigate potential relationships between instructor education and specific areas of language teaching believed to influence the SLA process. And finally, future studies should investigate instructor SLA education in addition to other instructor ID factors proven to have an affect on beliefs and teaching, such as instructor teaching experience, in the same study.

INSTRUCTOR TEACHING EXPERIENCE

Overview

Research investigating instructor teaching experience has been extensive in the field of teacher education (e.g., Carter, Cushing, Sabers, Stein, & Berliner, 1988; Fogarty, Wang, & Creek, 1983; Peterson & Comeaux, 1987) and has recently been extended to the field of SLA (see review in Borg, 2006a). In SLA, as in the general field of teacher education, the vast majority of these studies consist of separate discussions of novice/pre-service teachers (e.g., Almarza, 1996; Bigelow & Ranney, 2005; Borg, 2005; da Silva, 2005; Gutierrez Almarza, 1996; Johnson, 1992a; MacDonald, Badger, & White, 2001; Peacock, 2001; Pennington & Urmston, 1998; Richards, Ho, & Giblin, 1996; Sendan & Roberts, 1998; Urmston, 2003) or experienced/in-service L2 teachers (e.g. Andrews & McNeill, 2005; Borg, 1998b; Breen, 1991; Gatbonton, 1999; Johnston & Goettsch, 2000; Woods, 1996). Only a few studies have compared both novice/pre- and experienced/in-
service L2 instructors in the same study (Akyel, 1997; Johnson, 2003; Gatbonton, 2008; Mackey, et al., 2004; Mok, 1994; Nunan, 1992; Richards, 1998; Richards, Li, & Tang, 1998; Polio et al., 2006; Tsui, 2003); these comparison studies will be the focus of this section. An overview of these studies can be seen in Table 2 below.

*Table 2. Studies Comparing More and Less Experienced Language Instructors*

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Definition of Experience</th>
<th>Procedure</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akyel (1997)</td>
<td>5 experienced and 5 student ESL teachers in Turkey</td>
<td>Experienced teachers had an average of 11 years of experience and were seen as ‘competent’ by administrators &lt;br&gt;Student teachers were undergraduate education majors without teaching experience</td>
<td>Recording of one 45 min. lesson; SR protocol targeting the instructional goals, decisions, and related knowledge participants remembered considering while teaching</td>
<td>Experienced instructors considered more options in response to leaner cues in the classroom. Novice teachers focused on maintaining flow and were concerned re: choosing the appropriate strategy</td>
</tr>
<tr>
<td>Johnson (2003)</td>
<td>8 novice &amp; 8 expert language instructors in UK</td>
<td>Expert: ≥ 5 years of experience in task design &lt;br&gt;Novice: &lt; 5 years of experience in language teaching without task design</td>
<td>Participants asked to think aloud while responding to a task design assignment (average length 2.5 hours per participant)</td>
<td>Differences in task design abilities according to experience: experts demonstrated focus on one aspect of a task design at a time, sensitive to learners and the context of the task, and aware of the nuances of setting up and implementing tasks in-class</td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Experience Levels</td>
<td>Data Collection Methods</td>
<td>Findings</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Gatbonton (2008) | 4 novice and 7 experienced ESL instructors | Novice = 0-2 years of experience  
Experienced = ≥ 10 years of experience | Taped classroom observations, examination of transcripts of the lesson with researcher | Overall, novice and experienced instructors reported many of the same themes (classroom management, concern for student affective factors, etc.) although the frequency of each theme was different according to experience (though not at the level of significance) |
| Mackey et al. (2004) | 9 experienced and 9 inexperienced teachers of ESL | Experienced = had between 4 and 15 years of teaching experience (median = 10) and a masters in TESOL  
Inexperienced = undergrads with little to no formal teaching experience, enrolled in a TESOL methods class (no other TESOL education) | Each teacher taught one 30-minute lesson pre-planned by the researchers | Both groups used reactive implicit negative feedback most, although the experienced instructors provided much more reactive feedback compared to the inexperienced instructors.  
However, since the more experienced instructors also had master’s degrees in TESOL, the differences could not be solely attributed to experience |
| Mok (1994)       | 6 experienced and 6 inexperienced ESL teachers in US | Experienced = > 3 years  
Inexperienced: < 3 years | Journal entries, practicum reports and interview data | No significant difference in teacher reflections according to experience in any of the five categories: teacher role, attitudes towards learners, teaching strategies, materials used & expectations |
<table>
<thead>
<tr>
<th>Study</th>
<th>Number of Participants</th>
<th>Participants Characteristics</th>
<th>Data Collection Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nunan (1992)</td>
<td>9 ESL teachers in Australia</td>
<td>Teachers with a variety of experience, anywhere from new teachers to those who had been teaching 15+ years</td>
<td>1 lesson for each teacher was videotaped and observed. Each instructor then met with the researcher for a general SR</td>
<td>Less experienced teachers more concerned with classroom management; more experienced teachers with language issues</td>
</tr>
<tr>
<td>Richards (1998)</td>
<td>16 ESL teachers in Hong Kong</td>
<td>Experienced instructors had an average of 9.6 years of teaching and RSA English teaching certification and RSA diploma. Less experienced instructors had an average of 1.6 years of experience and a RSA certificate.</td>
<td>(1) Instructors completed a questionnaire about their teaching philosophy and lesson planning (2) Two classes were observed for each teacher (3) A follow-up interview with each instructor after each observed lesson</td>
<td>Both groups of instructors reported similar beliefs about the value of lesson plans. However, in practice, the two groups were different: more experienced teachers improvised more than inexperienced, using their lesson plans as outlines rather than step-by-step guides</td>
</tr>
<tr>
<td>Polio et al. (2006)</td>
<td>11 pre-service, 8 experienced ESL teachers; all of whom were NSs</td>
<td>Pre-service instructors had not formally taught and had ‘minimal’ interaction experience with NNSs. Half were enrolled in an ESL teaching methodology course. Experienced teachers had 4-27 years of experience. All had a master’s</td>
<td>Each instructor completed an information-gap task with an ESL learner; SR protocols with the researchers afterwards</td>
<td>No significant differences in # of recasts provided; significant difference in the way students used the feedback (more experienced teachers effective at encouraging learner output). More experienced instructors had strategies for getting learner-produced output and reported greater consideration for student learning.</td>
</tr>
</tbody>
</table>
degree in TESOL but one, who was in her final semester of the program.

| Tsui (2003) | 1 expert, 2 experienced and 1 novice ESL teacher in Hong Kong | Expert: in her 8th year of teaching  
Experienced: 5th year of teaching  
Novice: 2nd year  
All of these instructors were classified as “expert, experienced or novice” based on their teaching characteristics, not years of experience | Longitudinal design: class recordings, interviews, and examinations of instructors’ lesson plans and students’ work over 18 months  
In this study, the expert instructor’s behavior was characterized by her reflection, by her constant problem solving and searching for new ways to tackle the task at hand.  
The experienced instructors were more settled in their routines than the novice teacher, had found ways to encourage and reward participation, and were relatively consistent.  
The novice teacher acted reactively to most situations in the classroom, was inconsistent with her in-class behavior, and was still experimenting with her routine |
Operationalizing Instructor Teaching Experience

Before going into the specifics of the aforementioned studies, the issue of operationalizing instructor teaching experience must be discussed. Experience, like many other variables, is understood to be on a continuum though it is most often operationalized and investigated as a dichotomous variable. And, like many variables in SLA research, operationalizations abound when it comes to ‘more’ and ‘less’ experienced instructors, although the latter term has been much easier to define for obvious reasons.

On the less experienced end of the dichotomy, instructors are most often referred to as novice, inexperienced, or pre-service. Pre-service is the most unambiguous term and refers to instructors who have no formal teaching experience, most often those involved in language training and methodology coursework (Borg, 2006a). Inexperienced is rarely used and is often used synonymous with pre-service instructors (Akyel, 1997; Mackey et al., 2004). Novice is a term used in the literature investigating a language instructor of minimal years of teaching experience (Farrell, 2003; Pennington & Richards, 1997; Richards & Pennington, 1998; Spada & Massey, 1992) and is used in general expertise studies in the field of psychology (e.g., Dreyfus & Dreyfus, 1986), as well (discussed in more detail below). In teaching experience research, the term novice generally refers to instructors with little (generally less than 2 years) to no teaching experience in the L2 classroom (Farrell, 2003; Gatbonton, 2008; Pennington & Richards & 1997; Richards & Pennington, 1998; Tsui, 2003), however the term has been operationalized to include instructors with up to 5 years of experience (Johnson, 2003).

At the other end of the dichotomy instructors are referred to as experienced (e.g., Akyel, 1997; Gatbonton, 2008; Mackey et al., 2004; Polio et al., 2006; Richards, Tung,
An experienced instructor refers to an individual who has been teaching for quite some time. Studies vary greatly in their operationalization of experienced, describing instructors with as few as 3 years of experience (Mok, 1994) to more than 28 years (Polio et al., 2006). Several studies have zeroed in on the 7 year mark as qualifying instructors as experienced (Breen, Hird, Milton, Oliver, & Thwaite, 2001; Richards et al., 1992), with some researchers claiming that instructors’ experiences influence their beliefs and practices the most for the first seven years of their teaching, and slows down in subsequent years (Lopez, 1995, as cited in Berliner, 2001). The term in-service is more encompassing and refers to any instructor that is currently teaching, regardless of years of experience (see Borg, 2006a).

Expert, like novice, is a term used in general expertise literature, in research investigating teaching experience (e.g., Kagan, 1992; Calderhead & Shorrock, 1997; Turner-Bisset, 2001) as well as in a couple of studies on language teachers (Johnson, 2003; McNeill, 2005; Tsui, 2003). The determination of what qualifies an instructor as an ‘expert’ is an extremely complex notion, particularly when one is referring to a language instructor. Although expertise is beyond the scope of the current study, the term is important to discuss in order to highlight the differences between experience and expertise, which are at times used interchangeably and inappropriately so in the literature.

The notion of expertise originated from the field of psychology and initially focused on skills such as chess playing (e.g., de Groot, 1965). Researchers investigated the factors that went into reaching “expert” status, with a primary focus on the amount of experience, the procedural (automatized) knowledge gained in the field, and the minimal
effort required for retrieval (Glaser & Chi, 1988). Dreyfus and Dreyfus (1986) created a five-stage processing model that one normally goes through before becoming an expert.

The stages are: (1) novice; (2) advanced beginner; (3) competent; (4) proficient; and (5) expert. Dreyfus and Dreyfus (1986) are quick to point out that each person’s experience is unique and therefore everyone will not follow the progressive stages; some will bypass certain stages, some will spend more time in one or more stage(s) than others, and many will stop in a given stage and never reach expert status. Berliner (1992, 1995, 2001) extended this line of research to focus on expertise in teaching where it has been a heavily studied topic in the general field of teacher education, centering on how one becomes an expert instructor and what prohibits others from reaching the final stage. Berliner acknowledges the difficulty of defining expertise stating that, “although inexperience is equated perfectly with novice status in a field, the acquisition of experience does not automatically denote expertise” (2001, p. 466).

According to Berliner, the first two stages (the novice and advanced beginner) are characterized by conscious deliberation in action and thinking. Upon reaching the third stage, the competent teacher is characterized by taking personal responsibility for what occurs in the classroom. The fourth stage moves beyond rational thought during teaching into a reliance on the instructor’s intuition, and the expert stage is defined, in part, by the instructor’s ability to automatically and instinctively react and adapt in the classroom, without rational deliberation. Berliner (1995) summarized the main characteristics of the five stages stating,
If the novice is deliberate, the advanced beginner insightful, the competent performer rational, and the proficient performer intuitive, we might categorize the expert as often being arational (p. 4).

Berliner cautions that an instructor’s level of expertise is not a list of qualities or behaviors, but a combination of one’s experiences, education and one’s reflections upon them that makes some individuals progress from one level to the next. Teaching expertise is inherently difficult to operationalize and, unlike chess, which has concrete measurable outcomes (a winner and a loser), it is hard to pinpoint what makes one person a “better” teacher than another. Language teaching is seen as even more complex than general education, as there is an additional consideration of not only expertise in knowledge but the instructors’ skill in the target language, as well as their knowledge of how contextual factors and learner (and instructor) internal factors affect SLA (Borg, 2006b).

For these reasons, expertise has either been equated with experience (Andrews, 2007; Johnson, 2003) or considered as a combination of education and teaching experience (McNeill, 2005). In Andrew’s (2007) review on studies comparing novice and expert teachers’ language awareness (awareness of the underlying structures of the language being taught), he explicitly operationalizes expertise in association with years of teaching experience. Along with a list of general characteristics of the stage, Andrews states that teachers at the novice level include “all student teachers and 1st year teachers,” advanced beginners as “many second and 3rd year teachers,” competent instructors as “many 3rd and 4th year teachers,” equalizes proficient with “a modest number of teachers, from around the 5th year of teaching onwards,” and defines experts as “a small number of teachers, after at least 5 years” (p. 120). Johnson (2003) also equated years of experience
with expertise, defining ‘experts’ as instructors with “five or more years of experience” in task design, and labeling those without task design experience (but five years of experience in communicative language teaching in general) as “novices.” Tsui (2003) was the only study where both terms “experienced” and “expert” were used to differentiate between the two types of instructors. Tsui defined “experienced” instructors as instructors with 5 years of experience, while “expert” was reserved for an instructor with 8 years of experience and additional qualities (discussed in more detail below). Tsui explicitly differentiated between experience and expertise stating, "For some, eighteen years of experience is one year’s experience repeated seventeen times. Experience will only contribute to expertise if practitioners are capable of learning from it,” (2003, p. 13).

The reader is also reminded of McNeill’s (2005) study comparing NS and NNS instructors with and without expertise. McNeill operationalized expertise as a combination of both education (a master’s degree in TESOL) and teaching experience (at least two years). Not only does this operationalization collapse education and experience into one term – expertise – it cites and uses Berliner’s notion of expertise inappropriately.

The current study addresses the misuse and conflation of terms by investigating instructor education and years of experience separately. The operationalization of each of these terms used in the study will be defined in the Methods chapter. First, the studies comparing more and less experienced instructors will be critiqued.

Comparing More and Less Experienced L2 Instructors

Most studies have found differences according to instructors’ teaching experience. Akyel (1997) compared 10 instructors of ESL, 5 with an average of 11 years of
experience (referred to as “experienced”) and 5 student ESL instructors with no formal language teaching (although several had tutored English); all student instructors were education majors with an ESL minor. The instructors were additionally described as “competent” by school administrators. Akyel recorded one 45-minute lesson of each instructor and then met individually with each for a stimulated recall (SR) protocol targeting their instructional goals and decisions as well as any other related knowledge they considered while teaching. The SRs revealed that experienced instructors considered more options in response to learner cues in the classroom whereas novice teachers focused on maintaining flow and were concerned about choosing the appropriate management strategy or most effective technique.

Richards (1998) examined 16 ESL teachers in Hong Kong. The more experienced instructors had taught for an average of 9.6 years and possessed both a RSA certification and a diploma in English. The less experienced instructors had an average of 1.6 years of experience and only the RSA teaching certificate. The design consisted of a questionnaire targeting instructors’ teaching philosophy and beliefs about lesson planning, two recorded classroom lessons for each instructor, and two follow-up interviews with each participant (a follow-up interview after each recorded lesson). Although both groups of instructors were similar in their beliefs about lesson planning, in practice there were observable differences between more and less experienced instructors. Those with more experience improvised more than the inexperienced instructors, who were reluctant to deviate from lesson plans and used them as step-by-step guides.

Tsui (2003) is the only longitudinal study examining instructor experience to date. Over a period of 18 months, she studied 1 expert, 2 experienced and 1 novice ESL
teacher in Hong Kong. The expert instructor was in her 8\textsuperscript{th} year of teaching, the two experienced had taught for 5 years and the novice was in her second year. All of these instructors were classified as “expert, experienced or novice” based on their teaching characteristics, not years of experience. The design included class recordings, interviews, and examinations of instructors’ lesson plans and students’ work. Tsui found that, unlike what Dreyfus and Dreyfus (1986) would argue, in-class reflection and decision-making was a main characteristic of what it meant to be an “expert” teacher. In the study, the expert instructor’s behavior was characterized by reflection, by consistent problem-solving and searching for better ways to teach and explain concepts. The novice teacher, on the other hand, primarily acted in a reactive way in the classroom, was clearly still experimenting with her routine and, as a result, was inconsistent in her in-class behavior. The experienced instructors were more consistent than the novice instructor and were more settled in their routines. The experienced instructors had found ways to encourage and reward participation, although they were not as proactive as the expert teacher in continuing to reflect upon and adjust the lesson plans.

Johnson (2003) had a unique version of the term “expert.” She compared 8 novice and 8 expert language instructors in the UK, focusing specifically on task design. Instructors with at least 5 years of experience in task design were considered “experts;” those with less than 5 years were considered to be “novices,” although they each had more than five years of communicative language teaching experience in general. Participants were asked to think aloud while designing a task to the specifications of the assignment; the average think aloud lasted 2.5 hours per participant. Johnson articulated several differences in task design abilities according to experience: experts demonstrated
focus on one aspect of a task design at a time, were sensitive to learners and the context of the task, and were aware of the nuances of setting up and implementing tasks in class. The novices tended to skip from one aspect of the task to another, were less focused on the learners and the context where the task would be used, and did not focus on the pre- and between-task phase components. Johnson concluded the study by calling for more research focusing on specific aspects of language teaching.

In their (2006) study, Polio et al. compared native speaker (NS) opinions on videotaped NNS learner interaction data and found that NS perception varied according to teaching experience (any versus none). 11 pre-service, 8 experienced ESL teachers (all of whom were NSs) participated in the study. Pre-service instructors had not formally taught and had “minimal” interaction experience with NNSs; half were enrolled in an ESL teaching methodology course. The experienced teachers had a large range of teaching experience — 4 to 27 years — and all had a master’s degree in TESOL (except for one, who was in the last semester of her degree). Each instructor completed an information-gap task with an ESL learner and then met with the researchers for an SR protocol immediately following. Although there were no differences in the amount of recasts provided during the interactions according to instructor experience, there were significant differences in the amount of speech learners produced: more experienced teachers were more effective at encouraging learners to speak in class. The SRs revealed that more experienced instructors had strategies for encouraging learners to produce output and reported greater consideration for student learning, problems, and comprehension. Pre-service teachers were more concerned with their own performance, student feelings, and with the lesson plan and classroom management.
In the study most relevant to the current dissertation, Mackey et al. (2004) compared experienced and inexperienced ESL instructors’ use of focus on form (FonF) in a classroom lesson. The experienced instructors had an average of ten years of experience (range: 4-15 years) and all possessed a master’s degree in TESOL. The inexperienced instructors were undergraduate students with little to no formal teaching experience enrolled in a TESOL methods class (with no additional background in TESOL or SLA). Each taught one 30-minute lesson pre-planned by the researchers. Results reveal that both groups used reactive implicit negative feedback the most, although the experienced instructors provided much more reactive feedback in general compared to the inexperienced instructors. However, since the more experienced instructors also had master’s degrees in TESOL, the differences could not be solely attributed to experience.

A follow-up study examined the effect of a workshop on inexperienced instructors’ use of incidental FonF in the classroom (discussed in the previous section).

While many have found differences according to years of teaching experience, several studies have found no differences. Mok (1994) compared 6 experienced and 6 inexperienced instructors of ESL in the United States. “Experienced” was defined as greater than three years, while the term “inexperienced” was reserved for those instructors with less than three years of teaching. Mok asked instructors to journal on their teaching experiences and interviewed the instructors. The researcher found no significant difference in teacher reflections according to experience for any of the five categories that emerged from the data: teacher role, attitudes towards learners, teaching strategies, materials used and expectations. One must wonder if the minimal difference between the experiences of the two groups could explain the lack of differences. The
other issue that could have precluded any significant findings could be in the methods; without guidance on what to journal, instructors could have interpreted the task in several ways and thus various task interpretations could have mediated any affects of experience.

Gatbonton (2008) collected data from 4 novice ESL teachers and compared it to data collected earlier from 7 experienced instructors (Gatbonton, 1999). The researcher operationalized an instructor as “novice” if they had 2 years of experience or less; “experienced” instructors in the study had at least ten years of experience. In both studies, Gatbonton taped a random lesson, transcribed the lesson, and examined the transcript at an unspecified later date with each instructor. Overall, novice and experienced instructors reported many of the same themes (classroom management, concern for student affective factors, etc.), although the frequency of each theme was different according to experience. The experienced teachers were most concerned with language management, followed by procedure checks and progress reviews. The novice teachers were most concerned with student behavior and reactions, followed by language management and then procedure checks. However, these frequency differences were not significant.

**Limitations of Existing Instructor Experience Research**

In addition to the operationalization issues discussed prior to the studies, several other limitations must be addressed in future studies examining instructor experience. First, as is a main concern with the majority of research investigating the language instructor, studies examining instructor teaching experience have been greatly limited to one language: English, either as a second or foreign language. As previously discussed in the *Limitations* section reviewing NS/NNS instructors, English is an exceptionally
common language and not easily compared to other second or foreign languages. Research investigating instructor experience must be extended to other language settings to see if the differences observed in some studies are present regardless of context.

Second, research on this instructor ID has mostly focused on general teaching differences. Only Johnson (2003), Mackey et al. (2004), and Polio et al. (2006) examined specific features of language teaching. Rather than continuing to focus on overall teaching behavior, studies must narrow their focus to get at the heart of the differences between instructors. By thoroughly examining one topic (such as feedback provision) from many angles, this line of research will be able to begin to investigate if there are truly relationships between instructor IDs and teaching behavior critical to SLA, and what those relationships mean.

To this same end, researchers must expand their sample sizes to see if there are group differences. The largest study to date investigating instructor experience had eight participants per group (Richards, 1998); the others have had as few as five (Akyel, 1997).

With the exceptions of Polio et al. (2006) and Mackey et al. (2004), instructor teaching experience was the only ID examined in each study. As other variables such as NS/NNS status and instructor education have been found to influence instructor teaching behavior, researchers should analyze the same instructors according to multiple IDs to see which factors are most promising for this line of research.

Finally, in order to thoroughly understand the nature of differences according to language teachers’ IDs, researchers must not only examine differences in aspects of teaching behavior, but must also investigate why these differences exist. Studies must triangulate and collect both qualitative and quantitative data: not only classroom
observation and stimulated recall data, but belief questionnaire data, as well. These types of methods have often been employed separately. Studies investigating beliefs in relation to corresponding practices (discussed below) have shed light on the complex relationship and demonstrated the need for studies exploring if and how instructors’ beliefs influence aspects of their behavior believed to be important for SLA; namely, feedback provision.

Comparing Instructor Beliefs and Classroom Behavior

Overview

Instructor beliefs have been studied in isolation as well as in tandem with corresponding practices. Historically, most studies focused on beliefs about instructors’ general teaching practices (see Fang, 1996 for review of general education; for language teaching see Borg, 2006a), although researchers have increasingly focused on more specific aspects of language teaching, including grammar teaching (e.g., Andrews, 2003; Berry, 1997; Burgess & Etherington, 2002; Chandler, 1988; Chia, 2003; Eisenstein-Ebsworth & Schweers, 1997; Schulz, 1996, 2001), as well as literacy instruction (e.g., Beach, 1994; Davis, Konopak, & Readence, 1993; Grisham, 2000; Johnston, 2001; Olson & Singer, 1994, etc.). Of particular relevance for the current study, researchers have examined instructor beliefs regarding oral feedback provision (Árva & Medgyes, 2000; Basturkmen, Loewen, & Ellis, 2004; Chavez, 2006; Schulz, 1996, 2001), with few relating the beliefs to in-class provision (Árva & Medgyes, 2000; Basturkmen et al., 2004; Chavez, 2006).
Research Investigating Beliefs About Corrective Feedback

Schulz (1996) and (2001) investigated attitudes towards corrective feedback (within a study focused on grammar teaching in general) and compared instructors’ attitudes with students’. In her 1996 study, Schulz investigated the opinions of 92 teachers and 84 learners of foreign languages at an American university and found that learners reported corrective feedback as much more desirable than instructors. Only 42% of the instructors believed that spoken errors should be corrected while 90% of students said so. And 94% of the students disagreed with the statement “teachers should not correct students when they make errors in class” (my emphasis) as opposed to only 48% of the instructors. The findings were consistent regardless of the language studied/taught. Schulz replicated the study in 2001 and collected data from 122 FL teachers in Columbia along with 607 students and found results consistent with the original study. However, Schulz’s studies were restricted to beliefs and did not investigate whether or not they corresponded with actual error correction. Also, only two questions on the questionnaire targeted corrective feedback; the rest were about grammar teaching in general.

Studies Comparing Feedback Beliefs and In-Class Provision

Three studies have measured instructors’ beliefs about oral corrective feedback and subsequent in-class provision. In their 2000 study, Árva and Medgyes found associations between instructor beliefs and behavior although, as previously discussed, there were significant methodological issues with the design. In both datasets from the belief questionnaires (from Medgyes, 1994) and the classroom observation and interview data from the 2000 study, Árva and Medgyes found that NSs corrected less than NNSs
(although, again, no quantitative data were included in this statement). The researchers quoted a NNS instructor to explain these differences who said, “If natives don't speak the students' mother tongue, they cannot really ‘interpret’ the mistakes the students make” (p. 362). These attitudes about error correction were seen in the majority of NNS instructor behavior, including in areas such as lesson planning and overall course management.

Chavez (2006) also had a general focus in her study although, like Schultz, she also reported beliefs and behavior regarding feedback provision. In her semester-long case study of three instructors of intermediate-level FL German at an American university, Chavez utilized video recordings, interviews, student final grades and course evaluation to investigate instructors’ classroom language use and behavior in relation to their self-perceived role. Regarding feedback, instructors were internally very consistent in their own beliefs and behavior, although they differed considerably from each other. One instructor was primarily concerned with form, which was evident in her frequent explicit error correction. The second was primarily focused on making her students feel comfortable to talk and for that reason seldom corrected errors. The third and final instructor in the study was similar to the second; the instructor was concerned more with the class dynamic, flow and student participation than grammatical accuracy and this was reflected in the frequency of his use of elaboration, rather than explicit correction.

Unlike Chavez’s general focus, Basturkmen et al. (2004) specifically targeted one aspect of teaching: focus on form. In this study, the researchers defined focus on form (FonF) as “incidental time-outs taken by students and teachers to deal with issues of linguistic form during communicative lessons” (p. 243). Three male NS instructors of intermediate-level ESL were the focus of the case study, which consisted of four
observed/audio-recorded lessons, an in-depth interview, cued response scenarios and SR protocols. The SRs asked instructors what they were thinking at the time of the original interaction and what beliefs underlined these actions; instructors were also asked if they did what they thought they should have done in that situation.

The first instructor’s reluctance to correct errors was echoed in his behavior; both his beliefs and behavior favored student self-correction. The second reported an incongruent relationship between beliefs and practices of error correction, affirming that student self-correction was important while also discussing the importance of using recasts and other forms of correction. And while he also stated that he preferred to wait until the end of a lesson to discuss errors, the second instructor often interrupted the flow of communication to provide feedback. The third instructor had similar inconsistencies: while he believed it was important not to interrupt the flow of communication and that error correction should be unobtrusive, 44% of his feedback was complex (longer and tended to interrupt communication).

Results reveal that, when present, the differences were not contextually based as all taught the same type and level of language class. Rather, Basturkmen et al. related the differences to instructors’ individual beliefs and personal teaching style. The researchers closed with a call for studies investigating beliefs and behavior together, stating that future research on teachers’ beliefs, especially of unplanned events such as focus on form, need to be based both on beliefs and behavior (2004, p. 269).

Studies investigating instructor beliefs are of critical importance, as beliefs can influence practices in the classroom and therefore play a major role in determining the type of input and interaction students to which learners are exposed. Explorations into
potential relationships between teachers’ beliefs and classroom behavior are particularly crucial. As Chavez (2006) states, “The question of whether and how teachers’ self-perceived roles can account for variation in classroom-language use is of critical importance to all language instruction, but perhaps most to the FL setting, in which input is limited” (p. 52).

Limitations to Existing Belief/Practice Research on Feedback

First, the paucity of research investigating instructors’ beliefs in relation to feedback provision is an issue that must be addressed. In Borg’s (2006a) book-length review of language teacher cognition research, not one study specifically targeted teachers’ use of feedback. And while some studies reported error correction along with their general report of language teaching beliefs and behavior (Árva & Medgyes, 2000; Chavez, 2006), only study has specifically focused on instructor beliefs and error correction (Basturkmen et al., 2004).

All of the studies investigating instructor beliefs in relation to practice have had modest sample sizes. The Árva & Medgyes (2000) study had ten participants, while the others (Basturkmen et al., 2004; Chavez, 2006) were case studies of three instructors. Future studies should collect data from larger samples in order to test for statistical significance, which implies that the results could be found in other samples. Finally, the use of both qualitative and quantitative data (e.g., Basturkmen et al., 2004) should be encouraged, as it allows for more robust comparisons to other studies.
Rationale for the Current Study

The review of relevant literature highlights many areas in need for research. In contrast to the abundance of studies investigating learner individual difference (ID) factors that can and often do influence interactional feedback, there is a paucity of research focusing on IDs of the language instructor. Furthermore, the existing studies have limitations in scope and methodology that need to be addressed in future research.

Research comparing NS and NNS instructors has been limited to the ESL and EFL contexts, has employed very small sample sizes, and has failed to report or consider additional instructor ID factors (with few exceptions). Additionally, most studies have relied on a single measurement (often qualitative investigations of beliefs via questionnaires), and have often not investigated one specific teaching behavior.

The current dissertation addresses these limitations by extending NS/NNS instructor research to the Spanish foreign language classroom. By investigating multiple instructor ID factors, the current study does not assume that all NS instructors and all NNSs are the same. The study employs both qualitative and quantitative research methods and triangulates data collection to include participant background and belief questionnaires, videotaped classroom lessons, and stimulated recall methodology, employing the most robust research design to date targeting one aspect of language teaching. Finally, the current study enrolls the most participants thus far in this line of research. The maximum number of participants in any NS/NNS observational study has been 10 (Árva & Medgyes, 2000); the current study consisted of 60.

The existing studies in instructor education have been more concerned with instructor knowledge or measuring the changes in beliefs and behavior before and after
brief training courses. These studies have also focused on general teaching beliefs and practices and have thus far ignored corrective feedback.

This study addresses these limitations by investigating how instructor coursework in SLA is potentially related to beliefs and practices regarding a specific component of teaching: oral corrective feedback provision. Participants are not being assessed for changes during a course in which they are currently enrolled and thus the possible concern that their grade will be affected by what they report is eliminated.

For the variable of instructor experience, existing SLA research has employed very modest sample sizes and has been inconsistent in the operationalizations of what it means to be a “more” or “less” experienced instructor, often confounding the variable with expertise or combining it with education. Studies examining this variable have also, for the most part, been very general in nature, with only two investigating a specific component of language teaching.

The current study addresses these issues by employing the largest sample of participants to date (60) and taking great care in the operationalization of instructor experience by investigating instructor experience and education separately. It also focuses on a specific aspect of language teaching shown to influence SLA: feedback provision in the classroom. Following suggestions by Mackey et al. (2004) the study at hand utilizes stimulated recall protocols to examine why instructors differ in their feedback provision according to individual difference factors, and investigates how instructor IDs influence their feedback provision targeting specific linguistic targets in natural classroom settings.
As discussed in detail, instructor beliefs and practices have traditionally been examined in isolation. The studies that have investigated the relationship between beliefs and practices have been case studies or studies with ten or fewer participants, and have utilized qualitative measures exclusively.

The current study addresses these limitations by examining the beliefs and practices of a specific instructor behavior: in-class provision of oral corrective feedback. By triangulating with qualitative and quantitative methods, the current study has the most robust research design to date, and thoroughly investigates how instructor IDs influence their beliefs and behavior of oral corrective feedback provision in the classroom.

Taking into account the limitations and critical gaps in existing SLA research, the current study examines potential relationships between instructor ID factors and their: (1) beliefs about oral corrective feedback; (2) in-class provision of it; (3) the reasoning behind their in-class decisions regarding if, when and how to address learner errors with oral corrective feedback; and (4) explores the relationships between instructors’ beliefs and classroom behavior.
RESEARCH QUESTIONS

To this end, the research questions motivating the current dissertation were as follows:

RQ1: Do instructors of foreign language (FL) Spanish differ in their beliefs about the provision of oral corrective feedback in the classroom? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

RQ2: Do instructors of FL Spanish differ in the type and amount of oral corrective feedback they provide in the classroom? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

RQ3: Do instructors of FL Spanish differ in the consistency (i.e., providing $x$ type of oral corrective feedback for $y$ errors) of their oral corrective feedback provision? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

RQ4: Do instructors’ beliefs about how oral corrective feedback should be provided in the classroom correspond with how they actually provide feedback to learners? If not, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?
RQ5: Do instructors of FL Spanish differ in their reasoning behind their feedback provision choices? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

HYPOTHESES

Due to the lack of previous literature investigating the independent variables in the foreign language classroom, null hypotheses were tested.

HYP1: Instructors of FL Spanish do not differ in their beliefs about the provision of feedback in the classroom. Instructors of differing NS/NNS status, SLA education, or years of teaching experience do not differ in their beliefs.

HYP2: Instructors of FL Spanish do not differ in their provision of oral corrective feedback in the classroom. Instructors of differing NS/NNS status, SLA education, or years of teaching experience do not differ in their oral corrective feedback provision.

HYP3: Instructors of FL Spanish do not differ in the consistency of their oral corrective feedback provision. Instructors of differing NS/NNS status, SLA education, or years of teaching experience do not differ in the consistency of their oral corrective feedback provision.

HYP4: Instructors’ beliefs about how oral corrective feedback should be provided in the classroom do not correspond with their feedback provision. Instructors of differing
NS/NNS status, SLA education, or years of teaching experience do not differ in the relationship between beliefs and behavior.

HYP5: Instructors of FL Spanish do not differ in their reasoning behind their feedback provision choices. Instructors of NS/NNS status, SLA education, or years of teaching experience do not differ in their reasoning behind their feedback provision choices.
CHAPTER THREE
RESEARCH DESIGN AND METHODOLOGY

Introduction

The following sections discuss the research design and methodology employed in this study. The operationalizations and participants are described, followed by a thorough report of all aspects of the research design and the materials created for the project. Findings from pilot studies testing the research design and materials are also articulated in detail, as well as the changes that resulted from the three pilot studies. Finally, the chapter concludes with the coding and statistical procedures employed in the final version of the study.

Operationalizations

Prior to explaining the design of the study, constructs must be operationalized for the three independent variables (IVs). The researcher waited until all data had been collected in order to examine the participant information and appropriately operationalize and code for the IVs. This was motivated by the fact that this was the first study of its kind and there were no clear guidelines as to what constituted an “SLA educated” or “more experienced” Spanish FL instructor, particularly in terms of empirical studies. The operationalization decisions are explained in the following paragraphs.

Native/Nonnative Speaker Status

Instructor native speaker (NS) status was operationalized as having learned the language being taught (in this case, Spanish) as a primary, or one of several primary languages. This means that the instructor considered herself to be a native speaker and
was either (a) raised in a household where Spanish was used more than 50% of the time, or (b) attended elementary school where the primary language of instruction and communication was Spanish. This information was obtained from the participant instructors via the Background Questionnaire; explained in detail in the *Materials* section below. Using this operationalization, 38 native speaker (NS) and 22 nonnative speaker (NNS) instructors were identified in the dataset.

*SLA Education*

Waiting until all data had been collected to operationalize SLA education proved to be a prudent decision, as there were two main categories of participants: (a) those who had formally studied SLA theory or research, that is, having participated in several SLA courses; and (b) those who had taken only a teaching methodology course and no additional SLA courses, that is, participants required to take a methodology course but pursuing other fields of expertise outside applied linguistics. Thus, for the current study, SLA education was operationalized as having formally studied SLA theory or research. 20 instructors were classified as having SLA education and 40 had not.

*Teaching Experience*

Due to the imprecise and numerous ways “more” and “less” teaching experience has been operationalized, the researcher decided to investigate three operationalizations to see if one proved more able to distinguish between participants than another.⁶

---

⁶ The author thanks Melissa Baralt for her suggestion to use three operationalizations, and Rusan Chen for his assistance in ensuring that the three operationalizations were statistically comparable.
For the first operationalization, years of teaching experience were investigated as a continuous variable. In other words, the researcher simply used the number of years the instructor had been teaching (e.g., 3 years, 5 years, etc.). In this way, no categories or cutoffs were used and the researcher was thus able to examine patterns that emerged naturally from the dataset.

The second operationalization was based on the boundaries between each level of expertise (Andrews, 2007; Berliner, 1995). Although this study does not consider teaching expertise, the experience cutoffs in the model have been set forth because the literature implies there are instructor differences that emerge at these times of experience. For this reason, the cutoff years will be used only, without misappropriating the expertise terms such as “novice,” “expert,” etc. The reader is reminded that there are five levels in this model: the first level consists of instructors who have 0-1 years of teaching experience. The second level refers to instructors with 2-3 years of teaching experience. Instructors with 4 years of experience are next, followed by those who have been teaching 5 years. Finally, instructors with more than 5 years of experience are in the last level. Thus, this second operationalization defines teaching experience as a categorical variable with five levels.

The third and final operationalization was a dichotomy inspired by the methodology of existing literature on instructor experience (e.g., Johnson, 2003; Mok, 1994). The researcher used the cutoff of 7 years to refer to “more experienced” instructors and those with less than 7 as “less experienced” instructors, taking an average of the “more” and “less” operationalizations from the literature. Using this dichotomous operationalization, 25 instructors were identified as having more teaching experience and
35 were identified as being less experienced. And although 7 years was the proposed
cutoff, instructors in the study had either 6.5 years of experience or less, or 7 years or
more; there were none with experience between 6.5 and 7 years.

A correlation was run between the three experience constructs to see if the three
operationalizations would be equally representational ways of examining the data. A
correlation revealed that all three constructs were highly correlated (see Table 3) and thus
measured the same thing. Operationalizing these constructs is further discussed in the
*Coding Procedures* section.

*Table 3. Correlations for the Three Experience Operationalizations*

<table>
<thead>
<tr>
<th></th>
<th>Continuous</th>
<th>Categorical</th>
<th>Dichotomous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous Variable</td>
<td>1</td>
<td>.68***</td>
<td>.74***</td>
</tr>
<tr>
<td>Categorical Variable</td>
<td>.68***</td>
<td>1</td>
<td>.83***</td>
</tr>
<tr>
<td>Dichotomous</td>
<td>.74***</td>
<td>.83***</td>
<td>1</td>
</tr>
</tbody>
</table>

* *p < .05
** *p < .01
*** *p < .001

Description of Participants and Independent Variables

Participants were 60 instructors of university-level Spanish language courses
recruited from five universities in the Washington, DC area. Specifics on instructors’
native language(s), educational background, and teaching experience are reported in
Table 4.
### Table 4. Participant Background Information

<table>
<thead>
<tr>
<th>ID</th>
<th>L1</th>
<th>SLA Ed.?</th>
<th>Experience</th>
<th>Gender</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Spanish</td>
<td>Yes</td>
<td>6.5 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>2</td>
<td>English</td>
<td>Yes</td>
<td>3 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>3</td>
<td>English</td>
<td>Yes</td>
<td>3 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>4</td>
<td>English</td>
<td>No</td>
<td>5 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>5</td>
<td>Spanish</td>
<td>No</td>
<td>20 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>6</td>
<td>Spanish</td>
<td>No</td>
<td>10 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>7</td>
<td>Spanish</td>
<td>No</td>
<td>4 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>8</td>
<td>Spanish</td>
<td>No</td>
<td>9 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>9</td>
<td>English</td>
<td>No</td>
<td>2 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>10</td>
<td>English</td>
<td>Yes</td>
<td>30 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>11</td>
<td>English</td>
<td>No</td>
<td>40 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>12</td>
<td>Spanish</td>
<td>Yes</td>
<td>10 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>13</td>
<td>Spanish</td>
<td>No</td>
<td>6 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>14</td>
<td>Spanish</td>
<td>No</td>
<td>3 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>15</td>
<td>Spanish</td>
<td>No</td>
<td>30 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>16</td>
<td>Spanish</td>
<td>No</td>
<td>22 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>17</td>
<td>English</td>
<td>Yes</td>
<td>5 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>18</td>
<td>Spanish</td>
<td>No</td>
<td>1 month</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>19</td>
<td>English</td>
<td>No</td>
<td>3 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>20</td>
<td>English</td>
<td>Yes</td>
<td>4 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>21</td>
<td>Spanish</td>
<td>Yes</td>
<td>20 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td></td>
<td>Language</td>
<td>Learned?</td>
<td>Duration</td>
<td>Gender</td>
<td>Level</td>
</tr>
<tr>
<td>---</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>22</td>
<td>English</td>
<td>Yes</td>
<td>1 month</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>23</td>
<td>Spanish</td>
<td>No</td>
<td>4 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>24</td>
<td>Spanish</td>
<td>No</td>
<td>4 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>25</td>
<td>Spanish</td>
<td>No</td>
<td>28 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>26</td>
<td>Spanish</td>
<td>No</td>
<td>5 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>27</td>
<td>English</td>
<td>Yes</td>
<td>4 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>28</td>
<td>English</td>
<td>No</td>
<td>4 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>29</td>
<td>Spanish</td>
<td>Yes</td>
<td>8 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>30</td>
<td>English</td>
<td>Yes</td>
<td>8 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>31</td>
<td>Spanish</td>
<td>Yes</td>
<td>8 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>32</td>
<td>English</td>
<td>No</td>
<td>4 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>33</td>
<td>Spanish</td>
<td>No</td>
<td>28 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>34</td>
<td>English</td>
<td>Yes</td>
<td>4 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>35</td>
<td>English</td>
<td>No</td>
<td>6 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>36</td>
<td>Spanish</td>
<td>No</td>
<td>4 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>37</td>
<td>Spanish</td>
<td>No</td>
<td>20 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>38</td>
<td>Spanish</td>
<td>Yes</td>
<td>10 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>39</td>
<td>Spanish</td>
<td>Yes</td>
<td>4 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>40</td>
<td>English</td>
<td>No</td>
<td>3 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>41</td>
<td>Spanish</td>
<td>No</td>
<td>2 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>42</td>
<td>English</td>
<td>No</td>
<td>1 month</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>43</td>
<td>English</td>
<td>No</td>
<td>2 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>44</td>
<td>Spanish</td>
<td>No</td>
<td>8 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td></td>
<td>Language</td>
<td>Learn?</td>
<td>Years</td>
<td>Sex</td>
<td>Level</td>
</tr>
<tr>
<td>---</td>
<td>----------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>--------</td>
</tr>
<tr>
<td>45</td>
<td>English</td>
<td>No</td>
<td>3 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>46</td>
<td>Spanish</td>
<td>No</td>
<td>4 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>47</td>
<td>Spanish</td>
<td>No</td>
<td>1 month</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>48</td>
<td>Spanish</td>
<td>Yes</td>
<td>7 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>49</td>
<td>Spanish</td>
<td>Yes</td>
<td>2 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>50</td>
<td>Spanish</td>
<td>No</td>
<td>7 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>51</td>
<td>Spanish</td>
<td>No</td>
<td>12 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>52</td>
<td>English</td>
<td>No</td>
<td>35 years</td>
<td>M</td>
<td>Advanced</td>
</tr>
<tr>
<td>53</td>
<td>Spanish</td>
<td>No</td>
<td>39 years</td>
<td>M</td>
<td>Beginning</td>
</tr>
<tr>
<td>54</td>
<td>Spanish</td>
<td>Yes</td>
<td>4 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>55</td>
<td>Spanish</td>
<td>No</td>
<td>3 years</td>
<td>F</td>
<td>Intermediate</td>
</tr>
<tr>
<td>56</td>
<td>English</td>
<td>Yes</td>
<td>23 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>57</td>
<td>Spanish</td>
<td>No</td>
<td>25 years</td>
<td>F</td>
<td>Beginning</td>
</tr>
<tr>
<td>58</td>
<td>Spanish</td>
<td>No</td>
<td>3 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
<tr>
<td>59</td>
<td>Spanish</td>
<td>No</td>
<td>3 years</td>
<td>M</td>
<td>Intermediate</td>
</tr>
<tr>
<td>60</td>
<td>Spanish</td>
<td>Yes</td>
<td>8 years</td>
<td>F</td>
<td>Advanced</td>
</tr>
</tbody>
</table>

This demographic information was obtained via the Participant Background Questionnaire (see discussion in Materials section) and used to determine the operationalizations of the independent variables (see Operationalizations section).
Research Design

Data was collected in several ways in order to address the research questions posed in the current dissertation. All instructors had one 50-minute lesson videotaped and all completed the Participant Background Questionnaire and Participant Oral Corrective Feedback (OCF) Belief Questionnaire. Instructors who were willing and able to participated in a stimulated recall (SR) protocol following their classroom recording (see Stimulated Recall section for specifics on these participants); 35 of the 60 instructors participated in SR protocols. All components of the research procedure took the strengths and limitations of the existing research outlined in the literature review into consideration, and addressed the independent variables in a more robust way than has been done in previous studies. The study design can be seen in Figure 1; all aspects of the study are explained in detail in the following sections and were approved by the Institutional Review Boards from all five of the participating universities.

---

7 As some classes were longer, the first 50 minutes of each class were recorded and analyzed.
The fact that 35 of the 60 instructors participated in SR protocols before completing the Belief Questionnaire, while 25 did not, was originally a methodological concern. To ensure that participation in the SR protocols did not influence instructors’ responses, the researcher qualitatively compared the Belief Questionnaires from those who did and did not participate in the SR protocols. No differences were found between the two groups and thus this design decision was not considered a limitation.
Participant Recruitment

As previously articulated, participants in the study were 60 Spanish FL instructors recruited from five universities in the Washington, DC area. Any instructor teaching a Spanish language course (beginning, intermediate or advanced) during the summer or fall semester of 2009 was recruited and eligible to participate in the study. After receiving permission from the Language Program Directors from each university, the researcher recruited individual instructors via email during the summer and early fall of 2009. The email (see Appendix A) explained the study, invited instructors to participate, and contained links to the online Participant Consent Form (Appendix B) and the Participant Background Questionnaire (Appendix C) for the instructors to complete. Participants were not told the specific focus of the study until all aspects of data collection were completed; they were simply informed that the researcher was interested in teacher-student interaction in the Spanish FL classroom and were provided more details when their participation was completed. Instructors were not compensated for their time, but they were presented with the opportunity to learn the various research techniques utilized in the study, and the researcher also shared the results with the participants, as well.

---

8 Although data from 60 instructors was collected, three classroom recordings could not be used, two SRs and four belief questionnaires. Two of the classroom recordings were eliminated because (1) the lesson plan was changed, focusing on a movie or story discussion or (2) because one instructor said the presence of the video camera made her uncomfortable and she was not herself (nor would be if the researcher were to sit in the class). Two SR protocols were eliminated due to technical difficulties with the recordings (the tapes could only be heard using headphones, rendering the SRs impossible), and four belief questionnaires were not completed. Thus, data from all 60 participants were used, but not always in their entirety.
Recordings of Spanish Language Classes

After participants were recruited, the researcher corresponded with the instructors via email to (1) obtain their Spanish language course syllabi and (2) to determine a convenient day to videotape a lesson of the instructor. All video recordings took place on a day where the lesson targeted an aspect of grammar (i.e., a mode, a tense, a grammatical structure, etc.), or a difficult pair of prepositions.\(^9\) This decision was made based on results from the three pilot studies which demonstrated that, in classes targeting these linguistic items, substantially more feedback was available for analysis as compared to content or vocabulary-based lessons. Recording lessons of a grammar target allowed for comparisons of feedback provision; it would not have been valid to compare feedback provision in a lesson where a story or movie is discussed with feedback provided in a lesson targeting grammar. Analyzing lessons of a grammar target also allowed for more robust comparisons between language levels. The linguistic foci of the lessons recorded are reported in Table 5.

<table>
<thead>
<tr>
<th>ID</th>
<th>NS</th>
<th>SLA Ed.?</th>
<th>Experience</th>
<th>Course</th>
<th>Linguistic Foci</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Yes</td>
<td>6.5 years</td>
<td>ADV</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>BEG</td>
<td>Preterit and imperfect tenses</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>No</td>
<td>5 years</td>
<td>ADV</td>
<td>Subjunctive mood</td>
</tr>
</tbody>
</table>

\(^9\) While the majority of recorded classes focused on grammar, a couple focused on the differences between *por* and *para*. *Por* and *para* are two meaning for the English word “for” and are extremely difficult for students of Spanish to learn. For this reason, the prepositions are taught at every level of Spanish, from the very beginning through the most advanced levels (at least at the universities involved in the current study).
<table>
<thead>
<tr>
<th>No</th>
<th>Year</th>
<th>Study</th>
<th>Beg/Adv</th>
<th>Time</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>No</td>
<td>BEG</td>
<td>Prepositions (por and para)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>No</td>
<td>BEG</td>
<td>Preterit tense</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>No</td>
<td>INT</td>
<td>Commands, subjunctive</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>No</td>
<td>BEG</td>
<td>Reflexive/ reciprocal verbs</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>No</td>
<td>ADV</td>
<td>Preterit and imperfect tenses</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>30</td>
<td>Yes</td>
<td>ADV</td>
<td>Subjunctive mood</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>40</td>
<td>No</td>
<td>BEG</td>
<td>Verbs hay/estar/ser</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>Yes</td>
<td>BEG</td>
<td>Relative pronouns</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>No</td>
<td>INT</td>
<td>Indirect object pronouns</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>No</td>
<td>INT</td>
<td>Present perfect</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>30</td>
<td>No</td>
<td>ADV</td>
<td>Active and passive voice</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>22</td>
<td>No</td>
<td>INT</td>
<td>“If clauses,” hypotheticals</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>Yes</td>
<td>BEG</td>
<td>Daily verbs in 2nd, 3rd person</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>No</td>
<td>ADV</td>
<td>Ser and estar</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>No</td>
<td>BEG</td>
<td>Prepositions por and para</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>Yes</td>
<td>BEG</td>
<td>Prepositions por and para</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>Yes</td>
<td>BEG</td>
<td>Verbs like soler, deber, etc.</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>Yes</td>
<td>BEG</td>
<td>Question formation, estar</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>No</td>
<td>BEG</td>
<td>Prepositions por and para</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>No</td>
<td>BEG</td>
<td>Preterit tense</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>28</td>
<td>No</td>
<td>INT</td>
<td>Subjunctive mood</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>No</td>
<td>ADV</td>
<td>Relative pronoun clauses</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>No</td>
<td>ADV</td>
<td>Relative pronoun clauses</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>INT</td>
<td>Relative pronoun clauses</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>Yes</td>
<td>8 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>Yes</td>
<td>8 years</td>
<td>BEG</td>
<td>Impersonal and passive se</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>Yes</td>
<td>8 years</td>
<td>BEG</td>
<td>Impersonal and passive se</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>No</td>
<td>28 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>No</td>
<td>6 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>INT</td>
<td>Prepositions a and en</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>No</td>
<td>20 years</td>
<td>ADV</td>
<td>Relative pronoun clauses</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>Yes</td>
<td>10 years</td>
<td>ADV</td>
<td>Relative pronoun clauses</td>
</tr>
<tr>
<td>39</td>
<td>1</td>
<td>Yes</td>
<td>4 years</td>
<td>ADV</td>
<td>Relative pronoun clauses</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>INT</td>
<td>Prepositions a and en</td>
</tr>
<tr>
<td>41</td>
<td>1</td>
<td>No</td>
<td>2 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>No</td>
<td>1 month</td>
<td>BEG</td>
<td>Preterit tense</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>No</td>
<td>2 years</td>
<td>INT</td>
<td>Subjunctive and commands</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>No</td>
<td>8 years</td>
<td>INT</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>BEG</td>
<td>Preterit tense</td>
</tr>
<tr>
<td>46</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>INT</td>
<td>Subjunctive and commands</td>
</tr>
<tr>
<td>47</td>
<td>1</td>
<td>No</td>
<td>1 month</td>
<td>ADV</td>
<td>Subjunctive, prepositions</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>Yes</td>
<td>7 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>Yes</td>
<td>2 years</td>
<td>INT</td>
<td>Subjunctive mood</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>No</td>
<td>7 years</td>
<td>ADV</td>
<td>Subjunctive, prepositions</td>
</tr>
</tbody>
</table>
Classroom data was recorded via a Sony Handycam video camera on a tripod in the back of the classroom, which the researcher set-up prior to the lesson and then collected after class. To avoid any observer effect, the researcher was not present during the recordings. Because of the relatively large sample size, the fact that the researcher was not present and the fact that students were not alerted to the nature of the study (per IRB approval that did not require student consent), recording one 50-minute class of each instructor should still have been able to give a representation of the feedback tendencies of the instructor. Videotaping one lesson each of a large sample of instructors rather than several classes of fewer instructors was seen as a more robust design than previous studies and addressed the common limitation of small sample sizes. This potential limitation of taping one class was further addressed in follow-up questions during the stimulated recall (SR) protocols discussed below.
**Stimulated Recall Protocols**

In order to investigate instructors’ in-class decision-making regarding feedback provision, SR protocols were administered to 35 participants, allowing for sufficient data to run statistical analyses. Any participant willing and able to participated in a stimulated recall protocol with the researcher. See Table 6 for additional information on the instructors who participated in SR protocols.

**Table 6. Stimulated Recall Protocol Participants**

<table>
<thead>
<tr>
<th>ID</th>
<th>NS</th>
<th>Education</th>
<th>Experience</th>
<th>Course</th>
<th>SR Protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Yes</td>
<td>6.5 years</td>
<td>Advanced</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>Intermediate</td>
<td>3 hours after recording</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>No</td>
<td>5 years</td>
<td>Advanced</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>No</td>
<td>6 years</td>
<td>Intermediate</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>No</td>
<td>30 years</td>
<td>Advanced</td>
<td>22 hours after recording</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>Yes</td>
<td>5 years</td>
<td>Beginning</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>No</td>
<td>1 month</td>
<td>Advanced</td>
<td>40 hours after recording</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>Yes</td>
<td>20 years</td>
<td>Beginning</td>
<td>40 hours after recording</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>Yes</td>
<td>1 month</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-----------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>No</td>
<td>28 years</td>
<td>Intermediate</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>No</td>
<td>5 years</td>
<td>Advanced</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>Advanced</td>
<td>1 hour after recording*</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>Intermediate</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>Yes</td>
<td>8 years</td>
<td>Beginning</td>
<td>3 hours after recording</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>Intermediate</td>
<td>22 hours after recording</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>No</td>
<td>28 years</td>
<td>Intermediate</td>
<td>24 hours after recording*</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>Intermediate</td>
<td>40 hours after recording</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>No</td>
<td>6 years</td>
<td>Intermediate</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>Intermediate</td>
<td>1.5 hours after recording</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>No</td>
<td>20 years</td>
<td>Advanced</td>
<td>22 hours after recording</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>Yes</td>
<td>10 years</td>
<td>Advanced</td>
<td>19 hours after recording</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>No</td>
<td>1 month</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>No</td>
<td>2 years</td>
<td>Intermediate</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>No</td>
<td>8 years</td>
<td>Intermediate</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>Beginning</td>
<td>1 hour after recording</td>
</tr>
<tr>
<td>46</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>Intermediate</td>
<td>1.5 hours after recording</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>Yes</td>
<td>7 years</td>
<td>Intermediate</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>Yes</td>
<td>2 years</td>
<td>Intermediate</td>
<td>2 hours after recording</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>No</td>
<td>7 years</td>
<td>Advanced</td>
<td>1.5 hours after recording</td>
</tr>
<tr>
<td>52</td>
<td>0</td>
<td>No</td>
<td>35 years</td>
<td>Advanced</td>
<td>19 hours after recording</td>
</tr>
</tbody>
</table>

* Participant eliminated from data analyses
After the Spanish language class was recorded, the researcher watched the tape and, as soon as possible (within 48 hours after the end of class and, in all cases but one, before the instructor taught again), the researcher and the instructor met for the SR protocol in the researcher’s or instructor’s office (an average of 7.3 hours after the class). During the stimulated recall session, the researcher and instructor watched ten representative interaction episodes from the taped classroom interaction (student error - instructor response (or lack thereof) - student response (if any) on a MacBook Pro 17” computer screen through the Mac program iMovieHD. The researcher chose episodes containing student errors that were and were not addressed with oral corrective feedback, as well as a mixture of types of error and types of feedback as the foci of the SR protocols (explained more in the Materials section), all within the first half of the class as an additional control. The researcher played each interaction episode and stopped the recording to ask the instructors what they were thinking at that moment during the class, and about how they decided what errors to address with oral corrective feedback, as well as what went into their choice about how to provide it. The SR protocols allowed the researcher to investigate the reasoning behind instructors’ in-class feedback provision (RQ5), and all protocols were audio taped via the iPod Voice Memo application for further data analysis.

*Stimulated Recall Protocols: Methodological Considerations*

When targeting online decision-making there are two introspective methodological options: concurrent methods such as verbal think-alouds, which ask the participant to verbalize their thoughts usually during a written experimental task or interaction simultaneously as it is happening (e.g., Leow, 1998, 2000, for a discussion of
concurrent data elicitation procedures), and retrospective methods, such as SR protocols, which ask the participant to recall their thoughts usually after an oral/aural experimental task or interaction, often as soon as possible, and explain what they were thinking at the time of the original interaction (Gass & Mackey, 2000). For the current dissertation, as the interactions were actual lessons in the classroom and the participants were the teachers who were occupied with teaching, concurrent methods were not an option and retrospective methods had to be used. SR protocols are arguably the most robust and popular retrospective method used in SLA studies (see discussion in Gass & Mackey, 2000), and allowed the researcher to best investigate the research questions at hand. As with all research methods, SRs have inherent limitations in their design, which were considered when interpreting the findings. First, the methodology relies on participants’ ability to recall and report the thoughts they had during the original interaction. As with any retrospective method, there is a risk that participants will report their thoughts during the protocol, and not the thoughts they were having at the time of the interaction as required in order to be able to use the data reliably. Thus, there is the risk that what is reported is not what occurred at the time during the original interaction. There is also the risk that the participant will not remember what they were thinking during the interaction, a limitation that may be reduced by conducting the SRs as soon as possible after the interaction. Finally, there is the inherent limitation of both introspective methodologies: participants will fail to report what they were thinking. Despite the methodological limitations of SR protocols, they are still the most robust option to investigate instructors’ in-class decision making and thus the best option for the current study.
Questionnaire Dissemination

As explained above, the Participant Background Questionnaire was administered electronically during the summer and fall of 2009. The second and final questionnaire, the Participant Beliefs Questionnaire, was also sent to participants via a link in email, once their class had been recorded and the stimulated recall protocol had been administered (if applicable). Each questionnaire was accessed through a link sent to the instructors via email and instructors had the opportunity to complete it at their convenience. Two instructors requested paper copies of the questionnaires, which they completed and then returned to the researcher’s department mailbox.

Materials

Several instruments were created for the dissertation: a Participant Background Questionnaire, questions for the stimulated recall protocol, and a Participant Oral Corrective Feedback (OCF) Belief Questionnaire. Each instrument is discussed in detail in the following sections.

Participant Background Questionnaire

As previously articulated, the purpose of the Participant Background Questionnaire was to determine participant classification prior to data collection. The questionnaire consists of a combination of Likert-scale items and short answer questions, which provided the researcher with closed data for statistical comparisons as well as open-ended data for qualitative analyses. This also permitted participants the opportunity to express themselves in both structured and unstructured ways, allowing for robust data
collection. Questions were formulated to uncover: (1) instructors’ language background; (2) instructors’ SLA education; and (3) their teaching experience. Examples of each type of question are articulated below; the complete questionnaire can be seen in Appendix C. Participants were automatically taken to the Participant Background Questionnaire upon clicking “Done” in the Participant Consent form; they also had the option of completing the questionnaire at their convenience anytime online. Although the survey questions were written in English, participants were able and encouraged to respond to the open-ended questions in English or Spanish, per their preference.

Example of **language background questions**:

Closed question: Do you consider yourself to be a NS of Spanish? *Yes/No*

Open question: How would you define “native speaker”?

Example of **instructors’ SLA education questions**:

Closed question: Have you attended course(s) on second language acquisition research or theory? *Yes/No*

Open question: Please describe the course(s) in detail, including information on the topics covered, and information regarding where, when, and for how long you attended.

Example of questions about **instructors’ teaching experience**:

Closed question: What types of Spanish courses have you taught? Please check all that apply: *Beginning Spanish, Intermediate, Advanced, Linguistics, Literature*

Open question: How many years have you been teaching Spanish language courses?
The electronic version of the Participant Background Questionnaire was created using SurveyMonkey.com, a frequently used questionnaire software available on the web. SurveyMonkey allowed the researcher to keep all questionnaire data confidential in a password-protected computer file, and also provided the researcher with data storage and analysis tools. Instructors were able to return to the questionnaire at their convenience via a hyperlink and the software required answers to certain questions and prohibited participants from using ranking values more than once, thus allowing for the most robust, secure and convenient data collection and storage possible. Results from the pilot study demonstrate how necessary measures to ensure on-target responses were for the design of the current study (see *Pilot Study* section).

**Stimulated Recall Protocol Questions**

The stimulated recall sessions focused on instructors’ reasoning behind their feedback provision during class. Questions targeted instructors’ on-line decision-making, asking them what they were thinking at the time, if they noticed the error, what made them decide (or not) to address the error with corrective feedback, etc. The complete list of questions used in the stimulated recall protocols can be seen in Appendix D. An example of a closed and open question used in the SR protocols can be seen below:

**Closed question**: At that moment, did you notice the student error?

**Open question**: What made you decide (or not) to address the error with corrective feedback?
Approximately ten interaction episodes were chosen for each SR protocol, including student errors addressed and not addressed with oral corrective feedback, and including a wide range of error and feedback types present. The errors were chosen from the first half of the class as an additional control.\textsuperscript{10} At the end of the SR, the researcher asked instructors if the interaction episodes chosen for the SR protocol were representative of the lesson and their general feedback provision, and instructors were asked how their feedback provision in the observed class compared with their usual feedback provision. All instructors, with the exception of one,\textsuperscript{11} stated that the episodes were chosen to be representative of both the lesson and their feedback style in general.

The SR sessions were audio-recorded, allowing for later transcription and analysis. Please see the \textit{Coding Procedures} section for details on how the SR comments were coded and qualitatively analyzed. Participants were asked which language they would prefer to use during the SR protocols (English or Spanish), and the researcher used the language that the instructor was most comfortable using; 12 SRs were conducted in Spanish only, 18 in English, and in 5 protocols both languages were used.

\textit{Participant Belief Questionnaire}

The purpose of the Participant Oral Corrective Feedback Belief Questionnaire was to uncover instructors’ general beliefs about the importance of oral corrective

\textsuperscript{10} The researcher decided to investigate feedback at the beginning of class for several reasons: (1) regardless of the length of the class, the first 50 minutes could be analyzed; and (2) based on the pilot studies, more variance in topic was expected at the end of class, where some instructors would most likely have more of a teacher-centered summary of the lesson, while others would continue with interactive tasks until the course ended.

\textsuperscript{11} One instructor did state that the class was not normal due to the camera present and that her teaching would never be normal in front of a camera. In light of this information, this particular participant’s class recording was eliminated. Her information was still used for the belief questionnaire data (RQ1).
feedback in the SFL classroom, as well as their preferences for the type and amount of feedback they provide to their students during class. A sample question follows:

**Closed question:** Instructors have the responsibility to correct all mistakes uttered by students. (This was measured via a Likert scale ranking of 1-5, 5 being *strongly agree* and 1 being *strongly disagree*. Operationalizations were provided for each ranking.)

**Open question:** What is your overall opinion on the use of oral corrective feedback in the classroom?

The Belief Questionnaire was also adapted electronically via SurveyMonkey.com and can be seen in its entirety in Appendix E. The Belief Questionnaire targeted several topics: (1) instructors’ beliefs about how much feedback they provide; (2) their beliefs about the importance of feedback; (3) instructors’ beliefs about what type of feedback they provide; and (4) their beliefs about what type of errors they should correct. There were both closed and open questions for each topic, allowing for quantitative and qualitative analyses (see Table 7).

**Table 7. Belief Questionnaire Topics**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Closed Questions</th>
<th>Open Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of feedback</td>
<td>5, 21A</td>
<td>6</td>
</tr>
<tr>
<td>Importance of feedback</td>
<td>1B, 21B, 21D, 21E</td>
<td>25</td>
</tr>
<tr>
<td>Use of implicit feedback</td>
<td>13B</td>
<td>14</td>
</tr>
<tr>
<td>Use of explicit feedback</td>
<td>13A</td>
<td>14</td>
</tr>
</tbody>
</table>
Prior to the dissertation data collection, all materials were piloted at least once. The details of the three pilot studies, including the results and the changes that resulted from these studies, are outlined in the following sections.

Pilot Studies

Overview

The three pilot studies took place between 2007 and 2009 and resulted in few but important changes. The classroom recording component and a version of the Participant Belief Questionnaire were originally used in the researcher’s Qualifying Paper (QP) in the spring of 2007, in a semester project in a Research Methods course during the fall of 2007 (henceforth referred to as “Methods”), and extended and revised for the dissertation pilot in March of 2009 (referred to as “Pilot”). The stimulated recall questions were first employed in the Methods study and were minimally edited for the Pilot, based on the success of the questions from the first study. Thus, only the Background Questionnaire had yet to be tested prior to the third and final pilot. Details and results of each pilot study are explained in the following sections, and can be seen summarized in Table 8.
Table 8. Pilot Studies of the Research Design

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Course/Target</th>
<th>Design</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Qualifying Paper</strong></td>
<td>7 instructors: 4 NS, 3 NNS</td>
<td>Intermediate</td>
<td>1. Three lessons of each instructor videotaped 2. Belief Questionnaire from 7 observed instructors and 33 others in Department</td>
<td>1. NS instructors corrected 71% of errors; NNS instructors 53% 2. NSs used more explicit feedback than NNSs, who used more implicit feedback 3. NS and NNS beliefs differed 4. Both NS and NNSs’ beliefs often did not correspond with classroom behavior</td>
</tr>
<tr>
<td>Spring 2007</td>
<td></td>
<td>Lesson foci: (1) prepositions; (2) tense or mood; (3) vocabulary</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Research Methods Course</strong></td>
<td>2 instructors: 1 NS, 1 NNS</td>
<td>Beginning</td>
<td>1. One lesson of each instructor videotaped 2. SR protocols 24 hours after taped lesson 3. Belief Questionnaire from instructors (and their students re: their feedback preferences)</td>
<td>1. NS corrected 85% of errors; the NNS instructor 60% 2. Implicit feedback used more by both; NS 82% of the time and the NNS 70% 3. Both instructors noticed all errors; but NNS (with SLA education) reported more factors in her</td>
</tr>
<tr>
<td>Fall 2007</td>
<td></td>
<td>Lesson focus: Vocabulary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The dissertation pilot study (Pilot) is explained in detail below, and results are compared with results from earlier pilots of the design.

In order to test the practicality of the complete research design during the final Pilot, the revised Belief Questionnaire, SR protocols and the new instrument (the Participant Background Questionnaire), four volunteer pilot participants from the Department of Spanish and Portuguese completed all phases of the study: (1) consent
form and Background Questionnaire, (2) taped lesson, (3) stimulated recall protocols, and (4) Participant Belief Questionnaires. Six additional volunteers from Georgetown University’s German Department completed both questionnaires so that item analyses could be run. All questionnaires were given to pilot participants in Microsoft Word format, so that they would be able to insert comments while they completed the questionnaires. The Pilot data collection took place immediately following receipt of approval from the Georgetown University IRB approval in March of 2009, as no revisions were requested.

**Pilot Participants**

The four pilot participants in the Department of Spanish and Portuguese were chosen based on their teaching schedule in the fall of 2009 (three of the four did not teach language classes during the fall 2009 semester) and their willingness to participate in the project. The four final participants consisted of three NNS instructors (one male and two females) and one NS instructor (female). All four participants were graduate students between the ages of 27 and 35 and all had been enrolled in the program for 4-5 years. Two were applied linguistics Ph.D. candidates, one was a literature Ph.D. candidate, and one was a theoretical linguistics Ph.D. candidate. All had between 3 and 6 years of teaching experience.

The rationale behind the modest pilot size was twofold. First, almost all aspects of the dissertation design had been previously piloted and minor changes had already been made. Second, it was imperative to the integrity of the study that the participants not know the focus before the classroom recording and the researcher wanted to maximize
the pilot pool for the actual dissertation data collection and thus eliminate as few potential dissertation participants as possible.

_Pilot Study Results: Background Questionnaire_

The researcher recruited the dissertation pilot participants via email. Once the participants had agreed to participate in the study, they were emailed the Background Questionnaire in Word format. The German Department participants were asked to complete both questionnaires within a month and email them back to the researcher (they were also given the option of having paper copies dropped off in their mailboxes; all preferred an emailed version). The four Pilot participants from the Department of Spanish and Portuguese were asked to complete the questionnaire within two weeks, prior to the taped classroom lesson.

Having participants complete the questionnaire in the Word document proved to be an excellent choice for a first-time use of the questionnaire. Out of an original pool of 12 participants who returned the Background Questionnaire, two participants from the German Department had to be eliminated due to their lack of answers and off-target responses. This result further motivated the researcher to use an online questionnaire collection technique, where questions could not be skipped. Other than the two participants who were eliminated, no participant reported any problems or difficulties with the questions that were posed in the Background Questionnaire, and an earlier version of coding procedures proved able to distinguish the four participants from each other. Unlike the final version of the dissertation where there were two separate independent variables of SLA education and teaching experience, in the pilot study the
categories were collapsed into linguistic expertise. Two of the four participants were classified as having low linguistic expertise (pedagogical or theoretical knowledge) and two participants were classified as having high linguistic expertise (possessing both pedagogical and theoretical knowledge).12

Thus, out of the four cells of the Pilot study, there were two participants who were classified as NNS without linguistic expertise; one who was classified as NNS, with linguistic expertise; and one NS with linguistic expertise, allowing for a comparison of all cells. The Background Questionnaire, which was created for this particular dissertation, proved to be able to distinguish between participants that, at the surface, would appear to be very similar: similar age, length of time teaching, etc. Item analyses revealed that no items were highly correlated; thus, no items were testing the same topic and no changes were necessary for the Background Questionnaire.

Pilot Study Results: Recordings of Spanish Language Classes

The dissertation pilot extended the line of research initiated by the researcher’s Qualifying Paper (QP), Native and non-native instructors of Spanish: Beliefs and use of oral corrective feedback in the foreign language classroom, which took place during the spring semester of 2007. In the QP study, the researcher videotaped three lessons of seven intermediate-level Spanish instructors, four NS and three NNS instructors. One of the three days was a vocabulary lesson, one focused on a specific grammatical item (in

12 This independent variable of linguistic expertise was determined to be confounded during the dissertation proposal defense, as it contained both teaching experience (referred to as pedagogical knowledge) and SLA education (theoretical knowledge). For the final version of the dissertation, these variables were separated.
this study, two difficult prepositions\textsuperscript{13} were the focus of that particular lesson) and one day focused on a verb tense or mood. The researcher transcribed the interactions and coded for type of error and type of feedback. The results of the QP study revealed differences in NS and NNS instructor provision of feedback, and differences between beliefs and in-class behavior. While NNSs viewed oral feedback provision as one of their main roles, as instructors they provided feedback for only 53% of errors, whereas NS instructors addressed errors with feedback 71% of the time. NS and NNS instructors also differed in the types of feedback they provided for certain types of errors, and both NS and NNS beliefs and practices often did not coincide. Examinations of the correlations between NS and NNS instructor beliefs and practices reveal interesting trends.

In the QP, the majority of NS beliefs and practices did not appear to be related. The only correlations that had high $r$-values were the relationships between NS beliefs and practices that some errors are more important to correct than others. In terms of which errors are more important to correct than others, large $r$-values were present for the relationship between beliefs and provision of feedback for lexical, morphological, and phonological errors. NNS instructors were less consistent. Though NNS instructors strongly disagreed that instructors have the responsibility to correct all mistakes uttered by students, the correlation was almost a negative one ($r = -0.99$), meaning that beliefs and practices are inversely related; by addressing 53% percent of errors with corrective feedback NNSs’ high rate of feedback provision was enough to statistically contradict their beliefs that instructors do not have the responsibility to correct all student errors.

The relationship between beliefs and feedback provision was not explored in the

\textsuperscript{13} As previously articulated, Spanish has two words for the English word “for,” por and para, which proves extremely difficult for students whose L1 is English.
Research Methods study, which investigated learner feedback preferences and investigating instructor in-class decision-making regarding learner errors.

Like the earlier QP and Research Methods studies, once the four participants agreed to participate in the final Pilot, the researcher asked for their Spanish language course syllabi to determine the best dates for data collection. Since two of the four courses were advanced-level Spanish courses, one was intermediate-level and one was a beginning course, it was especially important for the researcher to find days to tape the lessons when there was as similar a linguistic target as possible. Due to logistical constraints, there were very few dates that both the researcher and the participants would be available and recording lessons where there was a grammatical focus (i.e., a tense, mood, or grammatical structure) was not possible. The dates selected for the data collection were subsequently chosen for their shared focus: a movie or reading discussion. The researcher felt that, though the lessons would not lend themselves to as much oral corrective feedback as would a grammatical lesson, they would be as similar as possible in terms of a common class goal, would minimize the difference between language levels and would allow the researcher to see which linguistic target (communicative or grammatical) would be the most robust for the final design.

The day of the lesson, the researcher set up a Sony Handycam camera on a tripod in the back of the room prior to class and then either stayed and observed the class (for three of the four lessons) or left the camera and returned after class (for one); one lesson could not be observed as the researcher and fourth participant taught at the same time.
The researcher transcribed all interaction and coded for error and feedback type (explained in detail in the Coding and Analysis section). Due to the limited sample size, statistical analyses were not run. Overall error correction trends can be seen in Table 9.

Table 9. Dissertation Pilot Study Results: Overall Error Correction

<table>
<thead>
<tr>
<th>Participant</th>
<th>NS/NNS Status</th>
<th>Expertise</th>
<th># Errors</th>
<th># Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NS</td>
<td>+</td>
<td>32</td>
<td>6 (18.8%)</td>
</tr>
<tr>
<td>2</td>
<td>NNS</td>
<td>-</td>
<td>44</td>
<td>13 (29.5%)</td>
</tr>
<tr>
<td>3</td>
<td>NNS</td>
<td>+</td>
<td>40</td>
<td>15 (37.5%)</td>
</tr>
<tr>
<td>4</td>
<td>NNS</td>
<td>-</td>
<td>51</td>
<td>8 (15.7%)</td>
</tr>
</tbody>
</table>

The percentage of errors addressed with oral corrective feedback in the dissertation pilot was much lower than in the Qualifying Paper, which is at least partly attributable to the meaning-focused movie or reading discussion. Unlike the Qualifying Paper, the majority of feedback provided in the dissertation pilot was explicit (see Table 10).

Table 10. Dissertation Pilot Study Results: Error and Feedback Types

<table>
<thead>
<tr>
<th>Participant</th>
<th># Errors</th>
<th># Corrected</th>
<th>Feedback Type(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Morphosyntax</td>
<td>26</td>
<td>4 (15.4%)</td>
</tr>
<tr>
<td></td>
<td>Lexis</td>
<td>5</td>
<td>2 (40%)</td>
</tr>
<tr>
<td></td>
<td>Phonology</td>
<td>1</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>Unsolicited L1 Use</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>32</td>
<td>6 (18.8%)</td>
</tr>
</tbody>
</table>

| 2           | Morphosyntax | 23         | 6 (26.1%)   | Explicit (5); Implicit (1) |
There were also some interesting patterns that were evident in the dissertation pilot regarding the relationship between the four participants’ beliefs and in-classroom practices. While there were no belief differences according to instructor expertise, there were several differences between the NS instructor and the NNS instructors, regardless of expertise. Like the QP study, the NS reported that correcting errors in the classroom was much less important than the three NNSs, and saw her role as much less feedback-related than the NNS instructors. Unlike the QP data, this NS instructor’s belief and classroom behavior (i.e., feedback provision) corresponded; the NS instructor had the second lowest error correction rate (18.8%). The NNS instructors, just like in the QP and the Research Methods study, reported that oral corrective feedback was one of their main roles and extremely important in the classroom, but did not correct many errors. In terms of type of correction, the two instructors with lower expertise reported that they used more explicit
feedback techniques, and the NNS instructor with higher expertise reported using a combination of both explicit and implicit techniques; the NS instructor did not answer this particular question. Both NNS instructors with lower expertise did in fact use more explicit techniques in the classroom, as did the NNS instructor with higher expertise, though all used a combination of both implicit and explicit techniques. And while all three female instructors reported that they consciously provided feedback in class (regardless of NS status or expertise), the one male NNS instructor reported that he did not think consciously about providing feedback in the classroom. Though there was only one male participant in the Research Methods study (a NS), he also reported that he did not provide feedback consciously, whereas the female (NNS) instructor in that study did. The data from two participants is not enough to draw a conclusion, but gender would be an interesting instructor variable to address in the future.

Based on the experience with the data collection in the pilot, the researcher decided that all taped class lessons for the dissertation would be unobserved. While there were no instances of the participants looking at the researcher, or asking about her presence, this decision was made for several reasons: (1) with the number of participants targeted for the dissertation data collection, it was not realistic for the researcher to be present for every lesson; (2) the researcher was then able to videotape more than one class at a time; (3) some of the data collection were done at other universities, where the participants were unfamiliar with the researcher; and (4) not all of the participants participated in the SR protocols. While the researcher needed to watch, code and analyze all of the videotaped lessons, she only needed to immediately watch the tapes of the participants who would participate in the protocols; the researcher could watch the other
taped lessons at a later date. The researcher would not have been able to observe only the
lessons being further used for SR protocols; this would differentiate the lessons used for
SR protocols from those not used, and the risk of an observer effect would have been
present in only some of the classes.

_Pilot Study Results: Stimulated Recall Protocols_

As previously mentioned, there were essentially two separate pilot studies for the
stimulated recall (SR) component of the proposed dissertation: one undertaken as a
semester project in a Research Methods course in the fall of 2007 (the “Methods” study)
and one conducted in the dissertation pilot in March of 2009 (the “Pilot”). The rationale
behind the SR component was to uncover _why_ NS and NNS instructors differ in their in-
class feedback provision.

In the Methods study, there were two participants (from an original pool of
four\(^\text{14}\)), one NS and one NNS instructor of beginning Spanish. Both had equal teaching
experience, and both participants had taken the same Spanish Teaching Methodology
course offered in the Department of Spanish & Portuguese at Georgetown. The only
differences were the NS/NNS status of the instructors and the amount of educational
background in applied linguistics.

Both classes were taped on the same day, when the target item was indefinite and
definite articles. The SR protocols were conducted 24 hours after the taped classroom
lessons (which were unobserved by the researcher), after the researcher had converted the

---

\(^{14}\) The Methods pilot study originally had four participants, two intensive beginning instructors and two intensive advanced instructors. However, one of the advanced instructors deviated from the lesson topic for the day, and had a conversation class instead of the indefinite and definite articles. Thus, both intensive advanced recordings were eliminated from the study.
mini DV tape to a standard DVD, which was played on the researcher’s laptop computer. In each SR, the researcher asked the instructor about ten interaction episodes, five of which contained errors addressed with feedback, and five which contained errors that were not addressed with oral corrective feedback. The first ten errors in the lesson were utilized for the SR protocols, since there were only eleven errors in one of the classes. Because the researcher converted the data to standard DVD, the time of each interaction episode was unknown and the researcher had to rely on instructors’ physical cues (i.e., walking to the board, etc.) as markers of where the interaction episodes were, which was difficult to follow and caused frequent fast-forwarding and rewinding of the DVDs.

Both the NS and NNS intensive beginning instructors noticed all the errors. Though they both noticed the errors, the female NNS instructor (with more SLA education) reported making extensive conscious online decisions regarding which errors she addressed with oral corrective feedback. Her decisions were primarily based on specific student characteristics, including individual student abilities, the context of the question, whether or not the error was a target item, if the error had been made more than once during the class period, if the mistake was likely due to a transfer error from their L1 or previously studied foreign languages, etc. The male NS instructor (with less SLA education) reported relying more on intuition and did not articulate the reasoning for choosing certain types of feedback. This does not necessarily mean that the NS did not have as much online reasoning as the NNS; he might just have not reported as much online processing. These potential differences of reporting are a potential limitation of the stimulated recall technique and are unavoidable, but important to keep in mind. One of the advanced-level instructors deviated from the linguistic target on the syllabus and held
a movie discussion instead, so the advanced NS and NNS instructor data was incomparable and subsequently eliminated from the SR part of the study.

For the second and final dissertation pilot conducted in March of 2009, the procedure was modified based on the early pilot experiences. This time, the researcher kept the mini DV tape as is so that time could be recorded and used to more efficiently locate each interaction episode. For the four classes that the researcher recorded, the following notes were taken for each error heard: (1) time on the recording tape; (2) student error; and (3) instructor response. The time on the recording tape proved exceptionally helpful, as the researcher was able to skip to the correct part of the tape and play each interaction episode without spending time searching for desired moment (which occurred during the Methods project) and without spending time editing and cutting the interaction episodes using time-consuming software. For each class, the researcher and the instructors met one hour after each class. This was done in order to keep the time constant between the class and the SR session for all four participants. The researcher met with each participant privately in her office, and the participant watched the Interaction episodes on a 40” television screen, which was connected to the camcorder. The most effective video camera was one that had a remote control, which allowed the participant to easily stop the tape whenever desired (though neither did).

Each protocol was audio recorded and transcribed by the researcher. Though instructors were given the option of using Spanish or English for the SR session, all chose to use mainly English, with occasional phrases and expressions expressed in Spanish.

The stimulated recall data was analyzed qualitatively, following the procedures outlined in *Coding Procedures*. The quantitative findings can be seen in Table 11.
Table 11. Dissertation Pilot Study SR Data: Overall Trends

<table>
<thead>
<tr>
<th>ID</th>
<th>NS?</th>
<th>Expertise?</th>
<th># Errors</th>
<th># Noticed</th>
<th>% Noticed</th>
<th># Decided Consciously</th>
<th>% Decided Consciously</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>High</td>
<td>10</td>
<td>6</td>
<td>60%</td>
<td>4</td>
<td>66.7%</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Low</td>
<td>11</td>
<td>11</td>
<td>100%</td>
<td>5</td>
<td>45.4%</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>High</td>
<td>9</td>
<td>9</td>
<td>100%</td>
<td>8</td>
<td>88.9%</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>Low</td>
<td>11</td>
<td>9</td>
<td>90.1%</td>
<td>2</td>
<td>22.2%</td>
</tr>
</tbody>
</table>

As seen in Table 11, the researcher calculated the number of errors present in the interaction episodes that were used for each SR protocol, the number/percentage of these errors that each participant reported noticing, as well as the number/percentage of noticed errors the participant reported thinking about consciously (i.e., deciding consciously to provide feedback or not). The researcher also coded the factors that each participant took into account in terms of whether or not to provide feedback, per the categories outlined in the Coding Procedures section.

For errors that were addressed with feedback, instructors reported taking into account very few internal factors: students’ lack of knowledge of the correct form (e.g., had not yet learned the present participle); and student’s previous language experience (e.g., L2 French causes a particular student to make certain errors often). Instructors cited many more external factors as justification for feedback provision, including: the number of turns a student had already spoken (i.e., the instructor felt the student had already spoken several times, so correction at that point would not hinder communication); the objective of activity not being met (i.e., for a meaning-based activity, the meaning was being lost; this was mentioned three times); a long pause in the student’s speech; a
student’s eye contact with the professor, interpreted as asking for help (this was mentioned three times); the error was one of the main objectives of the course (for intermediate-level Spanish there is a large focus on the subjunctive form); the error was a common mistake in Spanish (the instructor cited the importance of eliminating that frequent error; mentioned twice); and the fact that the entire class looked confused after a student’s mistake (mentioned twice).

For errors that were not addressed with feedback, instructors cited many internal factors as contributing to their decision including: an avoidance of making a big deal of an error (the instructor did not want to “put the student on the spot”); comparisons of individual students (i.e., this one has a harder time speaking than the other one who just spoke); learner anxiety (mentioned by every instructor, at least twice); a feeling of not wanting to “step on the toes” of the students (mentioned twice); an awareness of difference in status between students (one class had a professor from the School of Foreign Service as her student; another class had only one male who was older and the rest of the students were younger females); students’ personalities (an instructor said that she could not step on the loudest student because then the shy students would never talk). Instructors also reported not correcting a student when they shared something personal, if a student rarely spoke in class and spoke during the recorded lesson, or if an instructor perceived that a student had overall low confidence in the class. The NS instructor also cited the fact that students’ interlanguage speech was often very difficult for her as a NS to understand so sometimes she reported that she did not hear all the errors or had a hard time knowing what vowel a student was trying to produce.
Instructors cited many external factors as shaping their decisions not to provide oral corrective feedback: the error not being the focus of the activity (e.g., in a meaning-based activity, if the error did not interfere with the meaning then the instructor would not tend to correct it; mentioned by both instructors at least three times each); the timing of the error in the class (at the beginning of the class, instructors all said they wanted to encourage students to speak, so the errors were less likely to be addressed with feedback); if something extra occurred that was not part of the lesson plan (for example, in one class a debate unexpectedly arose about cell phone use), the instructor did not correct because she wanted to encourage students to speak; and one instructor cited the goal of the lesson as not targeting a certain grammatical structure and thus she did not feel the need to address it with feedback. Instructors also cited the frequency of a given student’s errors as a reason why they did not provide feedback (i.e., one student made errors every five words and the instructor felt it would be counterproductive to correct her); an avoidance of correcting a common error such as gender/number agreement that one instructor felt gradually eliminated itself over time; and one instructor said he often waited until the student was finished with their thought but often forgot to correct then or the error had passed too long ago to still be relevant.

Though there were only four participants, there were some interesting patterns from the limited data collected. The two participants with lower linguistic expertise reported fewer conscious feedback decisions in class (22.2% and 45.4%), whereas the two with higher linguistic expertise reported more (88.9% and 66.7%). The three NNSs reported noticing more errors (100%, 100% and 90.1%) than the NS instructor (60%). The instructor with the most teaching experience (10 years; also a NS) reported the
fewest conscious in-class decisions regarding student errors. While there was not enough
data to perform statistical analyses, the presence of trends at all in such a small sample is
promising and led support to the researcher’s final decision regarding the necessity of
having three independent variables.

There are two changes that resulted from the dissertation pilot study SR
component. First, while the large television screen allowed for easy presentation of the
Interaction episodes, the large television on a metal cart was cumbersome to use, and too
large to be used for the final data collection, especially since data was collected at five
different universities. For the dissertation itself, the researcher connected the video
camera to a MacBook Pro and played the video through IMovieHD application. The
second change that the researcher considered was in regards to the wording of the
stimulated recall questions. Originally in both earlier SR pilots, participants were asked
“Why do you think you noticed (or did not notice) the error at that time?” All participants
said they did not know and, as a result, the question was removed from the final version
of the questions. When the researcher specifically asked, “What made you decide (not) to
address that error with oral corrective feedback?,” three of the four participants asked the
researcher if that was the purpose of the study. The researcher decided not to change this
question for several reasons. First, since the instructor had already taped the lesson at the
time of the SR protocol, there was no potential influence of contaminating the lesson.
Second, the questions already began very generally (e.g., “What do you remember
thinking at that time?”), and became more specific (e.g., “Did you notice the student error
at the time?”). And third, since the Belief Questionnaire dealt mainly with feedback, if
the instructor did not discover the focus of the project during the SR protocol, she would
have once the final questionnaire was given. If all the questions had been changed to be more general in focus (e.g., “What made you decide to respond that way?”), by not asking specifically if the instructor noticed the error or not, the objectives of the SR protocols might not have been met. Following well-cited SLA studies that utilized SR protocols (e.g., Mackey, Gass & McDonough, 2000, etc.), no distracters (i.e., non-feedback-related interaction episodes) were used in the sessions.

Pilot Study Results: Belief Questionnaire Data

For the researcher’s Qualifying Paper (QP), a concise belief questionnaire consisting of 11 questions was given to both the seven instructors whose classes were recorded, as well as 33 other instructors in the Department. The questionnaire focused on the importance (or lack thereof) of providing oral corrective feedback in the classroom. Analyses revealed significant differences between NS and NNS instructors in terms of training in foreign or second language instruction; NNS instructors had significantly more training that discussed the use of oral feedback in the classroom than NSs.\textsuperscript{15} NS and NNS instructors also differed significantly in terms of the importance they placed on oral feedback in the classroom; NNSs reported that oral feedback in the classroom is more important than NSs. The final significant difference in beliefs of NS and NNS instructors was regarding the best time to correct errors: NSs reported the best time to correct errors is after the student is finished communicating (whether it be after a sentence or an extended dialogue), whereas NNSs reported that after the sentence containing the error is

\textsuperscript{15} It is important to point out that the question asked about training in language teaching and methodology; it did not explicitly consider SLA education.
the ideal time to correct errors. Overall, NNSs placed a higher importance on oral feedback in the classroom than NSs.

In the Research Methods study (fall 2007), the Belief Questionnaire was expanded to also ask instructors not only about the importance they placed on feedback provision in the classroom, but also if ID factors (specifically NS/NNS status and gender) influenced their feedback provision. Data from the four instructors completed the questionnaire in the Research Methods study (the two beginning instructors as well as the two advanced instructors whose classroom recordings and SR protocols could not be used) revealed results in keeping with the QP: the NNS instructor placed greater emphasis on provision of feedback as one of her major roles, while actually correcting less than the NS in practice. The NNS instructor believed that being a NNS affected her teaching while the NS did not believe his NS status affected his teaching choices. The NNS reported that having learned Spanish as a foreign language like the students allowed her to remember what types of feedback correction worked best for her, which errors were more likely to be static if not treated with feedback, and gave her a great sense of empathy for the students. While the NS instructor did not report that being a NS affecting his teaching of Spanish he did state, like the NNS, that learning a foreign language in the past made him think about his experiences learning a foreign language and how he best liked to be corrected. Though the intensive advanced classroom interaction data was not included in the analyses of the classroom interaction data and SR protocols, a quick analysis of the NS and NNS advanced instructors revealed the same beliefs as the beginning instructors. For both earlier pilots, a paper copy of the Beliefs Questionnaire
was given to instructors following the class recordings (Qualifying Paper) or after the SR protocols (Research Methods study).

For the dissertation pilot study, each instructor was emailed a copy of the Beliefs Questionnaire immediately following the SR protocol, and asked to return the questionnaire within a week. In this final Pilot, the Beliefs Questionnaire was greatly expanded to include the following topics: the importance of providing feedback in the classroom; instructor role in relation to feedback; the influence of instructor IDs on feedback provision; questions targeting the amount, type, timing and setting of feedback provision; the type of errors to correct; as well as factors affecting instructors’ in-class decisions regarding oral corrective feedback provision.

Analyses revealed that there were no belief differences according to instructor linguistic expertise, but there were several key differences between the NS instructor and the three NNS instructors, regardless of expertise. The three NNS instructors believed that providing oral corrective feedback in the classroom was much more important than the NS did, which was the same result found in the QP and the Research Methods study. Also similar to the previous two pilots, the NNS instructors saw their instructor role as much more feedback-related than the NS instructor. In terms of type of correction, the NNS instructor with greater expertise reported using a combination of implicit and explicit feedback techniques, the two instructors with less expertise reported that they used more explicit feedback techniques, and the NS instructor did not answer this particular question. And while all three female instructors reported that they consciously provided feedback in class (regardless of NS status or expertise), the one male NNS instructor reported that he did not think consciously about providing feedback in the
classroom. Six participants from the German Department also completed the Belief Questionnaire so that an item analysis could be run with data from ten participants.

Based on the item analysis, only two items on the Belief Questionnaire were highly correlated with each other, and thus one question was subsequently eliminated from the final version of the questionnaire (see final version in Appendix E).

Conclusions/Changes Resulting from the Pilot Studies

In conclusion, the dissertation pilot study demonstrated the robustness of the final research design. All instruments proved effective and efficient, and only minor changes were necessary: one question was eliminated from the Belief Questionnaire, the questionnaires were changed to electronic format, and the researcher decided to use non-observation recordings for the dissertation data collection. The results also revealed the necessity of dividing linguistic expertise into two separate variables – SLA education and years of teaching experience – in order to truly be able to examine instructor differences in beliefs and practices regarding in-class feedback provision. This was a particularly important decision for the coding and analyses that took place in the final study.

Coding Procedures

As this was the first study of its kind, bridging the gap between several fields and types of empirical studies, great care was taken to make sound coding decisions. The following section details the coding procedures that took place, and explains the rationale motivating each decision.
Participant Background Questionnaires

The purpose of the Participant Background Questionnaire was threefold: (1) to obtain demographic and background information regarding the participants; (2) to determine participant NS/NNS status; and (3) to determine the SLA education (+/-) and teaching experience (more or less) for each participant.

To code for instructor NS/NNS status, questions 24, 26 and 28 were examined. Instructors who reported themselves as being NSs (Q24) and spoke Spanish more than 50% of the time at home (Q26) or attended primary school where Spanish was one of the main languages spoken (Q28a) were considered native speakers for the study.

Instructors were coded as + SLA education if they reported having taken courses in second language acquisition theory or research (Q8 and Q9). All instructors who were classified as having SLA education had coincidentally taken two or more courses.

To code for instructor teaching experience, question 1 on the background questionnaire was examined. As previously explained, three operationalizations were calculated for each participant. The first considered teaching experience as a continuous variable and thus years of experience (participants’ answer to question 1) was used. The second operationalization investigated teaching experience as a categorical variable. An instructor was coded as ‘0’ if they had zero-one years of teaching experience and as ‘1’ if they had two to three years of experience. A coding of ‘2’ was assigned to instructors who had been teaching for four years, ‘3’ to those teaching for five years; and a ‘4’ was assigned to those who had taught more than five years. Finally, for the third operationalization, experience was a dichotomous variable; ‘0’ referring to an instructor with 0-6.5 years of teaching experience and ‘1’ referring to an instructor with 7 or more
years of experience. All additional questions on the Background Questionnaire were used to qualitatively explain results.

Classroom Interaction Data

The classroom interaction data was coded for type of error (lexis, morphosyntax, phonology, non-corresponding answer, and unsolicited use of L1) and type of feedback (explicit or implicit). Following Ellis (2008), explicit feedback included overt corrections, metalinguistic information and elicitations. Implicit feedback included recasts, confirmation checks, clarification requests, elaborations, and negotiations for meaning. Definitions and examples of each of these operationalizations can be seen in Tables 12 and 13; all example data are taken from the pilot studies.

Table 12. Examples of Coding Procedures for Error Type

<table>
<thead>
<tr>
<th>Type of Error</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexis</td>
<td>Use of an inappropriate word or a word that does not exist in the target language</td>
<td>Learner: El debe cambiar de casa.*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Translation: He should change house.*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error explanation: The student attempted to circumlocute around the correct verb for to move, mudarse. This sentence should be “El debe mudarse.”</td>
</tr>
</tbody>
</table>

Although the researcher recognizes that feedback explicitness is on a continuum rather than a dichotomy, for statistical purposes coding feedback as a dichotomous variable was necessary.
| Morphosyntax | Missing words or incorrect word order; gender/number agreement; incorrect verb conjugation; tense or mood | Learner: El debe **buscar otra** chica.  
Translation: He should look for *another girl.*  
Error explanation: In Spanish, when the object receiving the action of the verb is human, a personal a must be present between the verb and the receiver. This sentence should be “*El debe buscar a otra chica.*” |
| Phonology | Incorrect accent placement or a mispronunciation violating the rules of the target language (i.e., pronunciation of /h/ at the beginning of a Spanish word, etc.) | Learner: ¿Debo usar el **preterito**?  
Translation: Should I use the preterit*?  
Error explanation: The accent on the word preterit should be on the third to last syllable, not the penultimate. This question should be “¿Debo usar el preterito?” |
| Unsolicited L1 Use | Student-initiated use of the L1 (in most cases, English) | Learner: La **family** es muy importante en Cuba.  
Translation: Family* is very important in Cuba?  
Error explanation: The student used the English word for family instead of the Spanish word, familia. This sentence should be “La familia es muy importante en Cuba.” |
| Non-corresponding answer | An answer that does not correspond with the question | Instructor: ¿Qué es el sueño de tu vida?  
Learner: Mi madre |
Table 13. Examples of Coding Procedures for Feedback Type

<table>
<thead>
<tr>
<th>Type of Feedback</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overt Correction</strong></td>
<td>Immediate provisions of the correct answer, while simultaneously indicating an error was committed (Ellis, 2009)</td>
<td>Learner: Pienso que el capitán <em>crear</em>…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor: (interrupts) <em>crea</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Translation:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learner: I think that the captain to create*…</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor: creates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error explanation: The learner failed to conjugate the verb and used the infinitive form, <em>crear</em>.</td>
</tr>
<tr>
<td><strong>Metalinguistic Information</strong></td>
<td>Feedback providing the learner with specific metalinguistic information in response to an error (Lyster &amp; Ranta, 1997)</td>
<td>Instructor: ¿Cómo se dice ‘Gabriel García Márquez escribió <em>Cien Años de Soledad</em> en la voz pasiva?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learner: Márquez… (pauses)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor: Márquez es sujeto activo entonces no puede ser sujeto pasivo.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Translation:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor: How does one say, “Gabriel García Márquez wrote 100 Years of Solitude” in the passive voice?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor: Márquez is the active subject so he can’t be the passive subject.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error explanation: The learner had the</td>
</tr>
</tbody>
</table>
| **Elicitation**  
* (Explicit) | Incorrect passive voice word order by starting with the active subject.  
|---|---|
| **Elicitation**  
(Explicit) | Attempts by the instructor to directly elicit the correct construction from the learner  
(Lyster & Ranta, 1997) |
| Learner: Queremos que todos los niños *ir* a la escuela.  
Instructor: Entonces queremos que todos los niños...  
Learner: ¡Oh! Vayan.  
Translation:  
* Learner: We want all the children *to go* to school.  
Instructor: So we want all of the children...  
Learner: Oh! Go. (subjunctive).  
Error explanation: The learner failed to conjugate the verb using the subjunctive mood. |
| **Confirmation Check**  
(Implicit) | Utterances used to ensure that the interlocutor understood what the learner said  
(Long, 1983) |
| Instructor: ¿Qué quieres ser?  
Learner: Quiero ser *escritura*.*  
Instructor: ¿Quieres ser *escritura*?  
Translation:  
* Instructor: What do you want to be?  
S: I want to be *writing*.*  
Instructor: You want to be *writing*?  
Error explanation: The learner used the vocabulary word for “writing” in general, *escritura*, instead of the word for “writer,” *escritora*. |
| **Clarification Request**  
(Implicit) | Expressions “designed to elicit clarification of the learner's preceding utterance”  
(Long, 1983, p. 137) |
| Instructor: ¿Qué es censo?  
Learner: Es *una* record de gente.  
Instructor: ¿Es un qué?  
Translation: |
### Recast (Implicit)
Rephrasing a learners’ previous non target-like utterance “by changing one or more sentence components …while still referring to its central meaning” (Long, 1996, p. 434)

<table>
<thead>
<tr>
<th>Learner: Él asusta.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor: Él está asustado, sí.</td>
</tr>
<tr>
<td>Translation: Learner: He scares.*</td>
</tr>
<tr>
<td>Instructor: He is scared, yes.</td>
</tr>
</tbody>
</table>

**Error explanation:** The student did not use the conjugation to express a state of being scared.

### Negotiation (Implicit)
A mutual effort between the teacher and student to understand each other during a series of exchanges (Gass & Mackey, 2006)

<table>
<thead>
<tr>
<th>Learner: ¿El socio??</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor: El socio es el concepto. Esta persona persigue a los deudores, a las personas que tienen deudas para que estas personas paguen. Porque él, lo que quiere es recibir dinero. El quiere (gestures to himself as receiving something)…¿qué?</td>
</tr>
<tr>
<td>Learner: ¿Préstamos??</td>
</tr>
<tr>
<td>Instructor: No, él quiere...</td>
</tr>
<tr>
<td>Learner: (no answer)</td>
</tr>
<tr>
<td>Instructor: cobrar</td>
</tr>
</tbody>
</table>

**Translation:**
<table>
<thead>
<tr>
<th>Learner: The partner*?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor: The partner is the concept. This person pursues debtors, the people that have debts, so that these people pay. Because he, what he wants is to receive money. He</td>
</tr>
</tbody>
</table>
wants (gestures as receiving something)…what?

*Learner: Loans?*

*Instructor: No, he wants…to charge.*

Error explanation: The learner kept guessing the incorrect vocabulary word.

<table>
<thead>
<tr>
<th>Elaboration (Implicit)</th>
<th>Instructor provision of additional semantic information, often elaborating an original response with a paraphrase (Gass &amp; Mackey, 2006)</th>
<th><em>Learner: Se pobló.</em></th>
</tr>
</thead>
</table>
|                         | *Instructor: Se pobló, muy bien. O puedes decir “fue poblada” por los españoles antes que la ciudad de Jamestown.* | *Translation:*
|                         | *Learner: It was populated.*                                                                                   | *Instructor: It was populated, very good. Or you can say “it was populated” by the Spanish before the city of Jamestown.* |

Error explanation: The student did not use the target construction that was the purpose of the task.

A second researcher coded one-third of the error and feedback data for inter-rater reliability. The inter-rater was a fellow Ph.D. candidate in second language acquisition in the same department as the researcher. Inter-rater reliability measures (93% for type of error; 95% for feedback type) were high.

The data from these coded interactions allowed the researcher to investigate overall amount of feedback provided (% of errors addressed with oral corrective feedback), feedback preferences for type of error, and consistency of feedback provision (provided x type of feedback consistently for y type of error), as well as the relationship (or lack thereof) between the data and instructor NS/NNS status, SLA education, or years of teaching experience. This data assisted in addressing the second research question
(RQ2) (instructor provision of feedback), RQ3 (consistency of feedback provision), and part of RQ4 (investigating the relationship, or lack thereof, between instructor beliefs and practice).

**Coding of Stimulated Recall Protocols**

After the SR protocols were transcribed, the researcher counted the number of errors present in the interaction episodes targeted during the protocol, the number of errors reported as noticed by the participant, and the number of noticed errors the participant decided consciously to address or not address with oral corrective feedback. Feedback decisions were coded as ‘conscious’ if the instructors consciously thought about addressing or not addressing the student error with oral corrective feedback.

The researcher coded the factors that each participant took into account in terms of whether or not to provide feedback. The researcher categorized these factors as internal (i.e., learner affective factors, prior knowledge, etc.) or external (i.e., timing in the class, number of turns already taken, number of errors previously made, linguistic target of lesson, course objective, etc.) and further broke down factors as justification of why the instructor did or did not provide feedback. Thus, comments were organized in the four categories: internal factors cited for errors addressed with feedback; internal factors cited for errors not addressed with feedback; external factors cited for errors addressed with feedback; and external factors cited for errors not addressed with feedback. The same second researcher coded a third of all the SR data. Inter-rater reliability for noticed errors was 97%, reflected errors 95%, internal factors 95%, and external factors 94%.
Coding of Belief Questionnaires

As explained in the Materials section, the Participant Belief Questionnaire sought to uncover instructors’ beliefs regarding the following topics: (1) how much feedback they provide; (2) the importance of feedback; (3) what type of feedback they provide; and (4) what type of errors they should correct. There were both closed and open questions for each belief topic, allowing for quantitative and qualitative analyses (see Table 7, p. 115). The responses of each participant on the closed questions were grouped together according to topic and means were calculated so that quantitative analyses could be run (see Statistical Procedures below for specifics on which statistical analyses were employed). Participant responses for the open questions were also grouped together by topic and coded qualitatively to provide additional information on the closed questions, and to assist in explaining the results in relation to the three independent variables (e.g., if an instructor said they corrected a certain way due to their teaching experience, their NS/NNS status, or as a result of their SLA education or lack thereof). As with any true qualitative analysis, the researcher did not propose coding categories for the open questions, but rather looked for themes emergent in the data. Responses were compared by each of the three independent variables – NS status, SLA education, and experience – to see if any patterns were evident.

Statistical Procedures

To conclude this chapter, the following section outlines the statistical analyses used to address the research questions. Each research question is restated, followed by a description of the analyses employed.
RQ1: Do instructors of FL Spanish differ in their beliefs about the provision of oral corrective feedback in the classroom? If so, are these differences related to NS/NNS status, SLA education, or years of teaching experience?

For the first question, linear regressions were run due to the continuous nature of the dependent variable (the means of the belief topics). For each of the seven belief topics (amount of feedback, importance of feedback, implicit feedback, explicit feedback, lexical errors, morphosyntactic errors and phonological errors; see Table 7 on p. 115), means were calculated as the dependent variable and NS status, SLA education, and teaching experience were the three independent variables. One linear regression was run for each of the seven belief topics.

RQ2: Do instructors of FL Spanish differ in the type and amount of oral corrective feedback they provide in the classroom? If so, are these differences related to NS/NNS status, SLA education, or years of teaching experience?

The second research question required two types of statistical analyses: Chi-Square tests of independence for the first part of the question (type of feedback), as the dependent variables were categorical (implicit, explicit, or no feedback), and a linear regression for the second part of the question addressing the amount of feedback; the second part called for a linear regression due to the continuous nature of the dependent variable (percentage of errors addressed with feedback).

\[17\] It was necessary for the researcher to use means due to fact that not all participants answered each question and thus sums would not have been appropriate.
RQ3: Do instructors of FL Spanish differ in the **consistency** (i.e., providing x type of oral corrective feedback for y errors) of their oral corrective feedback provision? If so, are these differences related to NS/NNS status, SLA education, or years of teaching experience?

Chi-Square tests of independence were also required for this question, since the dependent variables were categorical (type of error and type of feedback). A Chi-Square test was run for each IV for each type of error: lexical, morphosyntactic, phonological, incorrect answer and unsolicited use of the L1, thus a total of 15 Chi-Square tests were run for this research question (e.g., one Chi-Square comparing NS and NNS instructors’ treatment of lexical errors, another comparing NS and NNS instructors’ treatment of morphological errors, etc.).

RQ4: Do instructors’ beliefs about how oral corrective feedback should be provided in the classroom **correspond** with how they actually provide feedback to learners? If so, are these differences related to NS/NNS status, SLA education, or years of teaching experience?

A correlation was needed to examine the potential relationship between instructors’ beliefs about a certain topic (e.g., amount of feedback that should be provided) and their corresponding practice in the classroom (i.e., the % of errors they address with oral corrective feedback), since both dependent variables (means of the belief topics and the measurement of in-class practice) were continuous (see Table 14).
Table 14. Belief Topic and Corresponding In-Class Data

<table>
<thead>
<tr>
<th>Belief Topic</th>
<th>In-Class Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of Feedback</td>
<td>Overall % of errors addressed with feedback</td>
</tr>
<tr>
<td>Importance of Feedback</td>
<td>Overall % of errors addressed with feedback</td>
</tr>
<tr>
<td>Use of Implicit Feedback</td>
<td>% of errors addressed with implicit feedback</td>
</tr>
<tr>
<td>Use of Explicit Feedback</td>
<td>% of errors addressed with explicit feedback</td>
</tr>
<tr>
<td>Treatment of Lexical Errors</td>
<td>% of lexical errors addressed with feedback</td>
</tr>
<tr>
<td>Treatment of Morphosyntactic Errors</td>
<td>% of morph. errors addressed with feedback</td>
</tr>
<tr>
<td>Treatment of Phonological Errors</td>
<td>% of phon. errors addressed with feedback</td>
</tr>
</tbody>
</table>

There were three independent variables: NS/NNS status, SLA education, and years of teaching experience. Data from the open-ended questions on the Belief Questionnaire were analyzed qualitatively and used to further explain the quantitative results.

RQ5: Do instructors of FL Spanish differ in their reasoning behind their feedback provision choices? If so, are these differences related to NS status, education, or years of teaching experience?

To answer the fifth and final research question, the stimulated recall transcriptions were coded and analyzed for: (1) percentage of errors reported as noticed, (2) percentage of noticed errors consciously decided to be addressed (or not) with feedback. Linear regressions were run comparing the percentage of errors addressed as noticed and the percentage of noticed errors reflected upon according to each IV (NS/NNS status, SLA education and experience). SR comments were analyzed qualitatively for (1) internal and
external factors taken into account when deciding to address an error with feedback and
(2) factors taken into account when deciding not to address an error with oral corrective
feedback.
CHAPTER FOUR

RESULTS

This chapter discusses the quantitative and qualitative results for the five research questions. For each question, the three instructor individual difference variables of NS/NNS status, SLA education, and teaching experience were investigated separately. That is to say, NS instructors were compared with NNS instructors, those with SLA education were compared to those without, and those with more teaching experience were statistically compared to those with less. This was done for two reasons: First, as this is the first study of its kind, the purpose was to see if one or more of these three variables were related to teachers’ beliefs, practices, and the reasoning behind their practices and thus independent analyses of each independent variable (IV) was most appropriate. The second reason for this statistical decision was based on the unequal number of participants in each of the eight potential groups (see Table 15). Though the researcher recruited heavily from five universities, the 60 participants willing to participate in the study were not equally divided in terms of their background and thus any investigation of interactions between variables was not statistically possible.

Table 15. Participants Organized by All Three Independent Variables

<table>
<thead>
<tr>
<th>NS, + SLA education, + experience</th>
<th>NNS, + SLA education, + experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 participants (8.3%)</td>
<td>3 participants (5%)</td>
</tr>
<tr>
<td>NS, + SLA education, - experience</td>
<td>NNS, + SLA education, - experience</td>
</tr>
<tr>
<td>4 participants (6.7%)</td>
<td>8 participants (13.3%)</td>
</tr>
<tr>
<td>NS, - SLA education, + experience</td>
<td>NNS, - SLA education, + experience</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>12 participants (20%)</td>
<td>2 participants (3.3%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NS, - SLA education, - experience</th>
<th>NNS, - SLA education, - experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 participants (28.3%)</td>
<td>9 participants (15%)</td>
</tr>
</tbody>
</table>

**Research Question 1: Beliefs**

To investigate if there were any differences in reported beliefs according to any of the three independent variables linear regressions were run due to the continuous nature of the dependent variable (the means of each belief topic); the independent variables were NS/NNS status, SLA education and teaching experience. One linear regression was run for each of the seven belief topics. Means and standard deviations from the seven linear regressions run for the belief questionnaire topics are reported in Table 16.
Table 16. Belief Topic Means and Standard Deviations

<table>
<thead>
<tr>
<th>NS/NNS Status</th>
<th>Education</th>
<th>Years of Teaching Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>NNS</td>
</tr>
<tr>
<td>Amount</td>
<td>1.76(.46)</td>
<td>1.89(.60)</td>
</tr>
<tr>
<td>Importance</td>
<td>3.47(.64)</td>
<td>3.64(.69)</td>
</tr>
<tr>
<td>Implicit</td>
<td>2.31(.76)</td>
<td>2.40(.88)</td>
</tr>
<tr>
<td>Explicit</td>
<td>1.61(.83)</td>
<td>1.54(.84)</td>
</tr>
<tr>
<td>Lexical</td>
<td>2.36(1.0)</td>
<td>2.81(.98)</td>
</tr>
<tr>
<td>Morphosyntactic</td>
<td>3.02(.66)</td>
<td>2.80(.85)</td>
</tr>
<tr>
<td>Phonological</td>
<td>1.62(.90)</td>
<td>1.35(.59)</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001
Amount of Feedback

Results reveal no significant differences according to any of the three independent variables for the belief topic amount of feedback. Regardless of NS/NNS status ($p = .33$), SLA education ($p = .64$), or years of experience ($p = .50$), instructors in the study reported that they believed it was important to be cautious with the amount of feedback provision during class time, stating that correction should only take place with the target item or if it impedes communication. Reported mean scores revealed a difference of less than .13 between each of the IV dichotomies (e.g., NS compared to NNS instructors, etc.), as reported in Table 17.

Table 17. Belief Questionnaire Topic 1: Amount of Feedback

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>-.15</td>
<td>-.98</td>
<td>.33</td>
</tr>
<tr>
<td>SLA Education</td>
<td>-.08</td>
<td>-.48</td>
<td>.64</td>
</tr>
<tr>
<td>Experience</td>
<td>.01</td>
<td>-.69</td>
<td>.50</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$

Importance of Feedback

Insignificant differences were also found for the belief topic importance of feedback. Regardless of NS/NNS status ($p = .24$), SLA education ($p = .17$), or experience ($p = .58$), instructors in the study all reported similar beliefs regarding the importance of providing oral corrective feedback (differences between the group means less than .23)
(see Table 18); all instructors agreed that feedback provision in the classroom was important.

Table 18. Belief Questionnaire Topic 2: Importance of Feedback

<table>
<thead>
<tr>
<th>NS/NNS Status</th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLA Education</td>
<td>-.27</td>
<td>-.139</td>
<td>.17</td>
</tr>
<tr>
<td>Experience</td>
<td>.01</td>
<td>.56</td>
<td>.58</td>
</tr>
</tbody>
</table>

* * p < .05
** p < .01
*** p < .001

 Implicit Feedback

With respect to opinions regarding implicit feedback, participants also did not differ. Instructors reported similar beliefs about the use of implicit feedback (differences of means less than .11), regardless of NS/NNS status (p = .60), SLA education (p = .61), or experience (p = .65) (see Table 19). In terms of instructors’ reported usage of implicit, explicit, and a combination of the two types of feedback, results indicate that instructors in this study believe that their use of implicit feedback falls between their use of explicit feedback (the least frequent type of feedback instructors believe they use) and a combination of explicit and implicit feedback.
Explicit Feedback

The fourth topic examined instructors’ beliefs regarding the use of explicit feedback. Like the previous topics, no significant differences were found for instructor reported beliefs about their use of explicit feedback in relation to NS/NNS status ($p = .80$), SLA education ($p = .42$), or experience ($p = .80$). Differences in means were less than .20 (Table 20); for the most part instructors believed that in the classroom, they used explicit feedback less than implicit and less than a combination of both types of feedback.

**Table 19. Belief Questionnaire Topic 3: Implicit Feedback**

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>- .13</td>
<td>-.52</td>
<td>.60</td>
</tr>
<tr>
<td>SLA Education</td>
<td>- .14</td>
<td>-.52</td>
<td>.61</td>
</tr>
<tr>
<td>Experience</td>
<td>.01</td>
<td>.46</td>
<td>.65</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$

**Table 20. Belief Questionnaire Topic 4: Explicit Feedback**

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>.06</td>
<td>.25</td>
<td>.80</td>
</tr>
<tr>
<td>SLA Education</td>
<td>.22</td>
<td>.82</td>
<td>.42</td>
</tr>
<tr>
<td>Experience</td>
<td>.00</td>
<td>.25</td>
<td>.80</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$
Lexical Errors

Turning to instructor beliefs regarding the correction of lexical errors, no significant differences were found for reported beliefs about the treatment of lexical errors and NS/NNS status ($p = .14$), SLA education ($p = .93$) or experience ($p = .56$). Differences in means were all less than .45 (see Table 21). For this question, instructors tended to report that lexical errors were more important to correct than phonological errors, but that morphosyntactic errors warranted the most feedback.

Table 21. Belief Questionnaire Topic 5: Treatment of Lexical Errors

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>-.46</td>
<td>-1.50</td>
<td>.14</td>
</tr>
<tr>
<td>SLA Education</td>
<td>-.03</td>
<td>-.08</td>
<td>.93</td>
</tr>
<tr>
<td>Experience</td>
<td>.01</td>
<td>.59</td>
<td>.56</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$

Morphosyntactic Errors

This belief that morphosyntactic errors are the most important to correct was common for all instructors and any differences were insignificant. Regardless of NS/NNS status ($p = .36$), SLA education ($p = .68$), or experience ($p = .92$), instructors all reported similar beliefs regarding the importance of addressing morphosyntactic errors with corrective feedback (differences between the group means less than .22) (see Table 22).
Table 22. Belief Questionnaire Topic 6: Treatment of Morphosyntactic Errors

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>.20</td>
<td>.93</td>
<td>.36</td>
</tr>
<tr>
<td>SLA Education</td>
<td>-.09</td>
<td>-.41</td>
<td>.68</td>
</tr>
<tr>
<td>Experience</td>
<td>.00</td>
<td>.10</td>
<td>.92</td>
</tr>
</tbody>
</table>

* *p* < .05  
** *p* < .01  
*** *p* < .001

Phonological Errors

Finally, instructors also did not differ in their opinions regarding phonological errors, reporting similar beliefs (differences of means less than .27), regardless of NS/NNS status (*p* = .28), SLA education (*p* = .96), or experience (*p* = .97) (see Table 23). As previously stated, instructors believed that phonological errors were the least important errors to address with corrective feedback.

Table 23. Belief Questionnaire Topic 7: Treatment of Phonological Errors

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>.26</td>
<td>1.10</td>
<td>.28</td>
</tr>
<tr>
<td>SLA Education</td>
<td>-.01</td>
<td>-.05</td>
<td>.96</td>
</tr>
<tr>
<td>Experience</td>
<td>-.00</td>
<td>-.04</td>
<td>.97</td>
</tr>
</tbody>
</table>

* *p* < .05  
** *p* < .01  
*** *p* < .001
**Research Question 1: Summary**

In summary, for the first research question, there were no significant differences in beliefs according to the three independent variables of NS/NNS status, SLA education, or years of teaching experience.

**Research Question 2: In-Class Feedback Provision**

The second research question investigated instructors’ feedback provision in terms of the type (implicit, explicit, or no feedback) and amount (percentage of errors addressed with oral corrective feedback) of feedback provided to learners in the Spanish FL classroom.

**Amount of Feedback**

A linear regression exploring the potential relationship between instructor NS/NNS status, SLA education, and experience (the independent variables) and amount of feedback (the dependent variable) demonstrated that only NS/NNS status is significantly related to the percentage of errors addressed with oral corrective feedback ($p < .01$). NS instructors corrected on average 74.7% of errors, 24.7% more than NNS instructors (ignoring SLA education and experience). There were also differences according to the amount of feedback provided in relation to education (an average difference of 11.6%) and in relation to experience (an average difference of 16.1%), however these differences were not significant ($p = .16$ and .08, respectively). Thus, the linear regression revealed significant differences in the overall amount of feedback provided (percentage of errors addressed with oral corrective feedback) according to
NS/NNS status (NS instructors corrected significantly more than NNSs), but not in relation to instructor SLA education, nor years of teaching experience (see Table 24).

Table 24. Amount of Feedback and Instructor ID Variables

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS/NNS Status</td>
<td>23.36</td>
<td>4.02</td>
<td>.00*</td>
</tr>
<tr>
<td>SLA Education</td>
<td>-7.04</td>
<td>-1.21</td>
<td>.23</td>
</tr>
<tr>
<td>Experience</td>
<td>.30</td>
<td>1.14</td>
<td>.26</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001

These trends are further demonstrated by error type in Figure 2 below.

Figure 2. Amount of Feedback
Feedback Type

In terms of type of feedback provided, Chi-Square tests revealed that NS instructors addressed errors with explicit feedback significantly more than NNSs (ignoring SLA education or experience): $X^2(2) = 102.63, p < .001$. Those without SLA education (ignoring NS/NNS status or experience) corrected with significantly more explicit feedback ($X^2(2) = 12.01, p < .001$) and those with less experience ignored errors significantly more than those with more teaching experience ($X^2(2) = 72.25, p < .001$).

Research Question 2: Summary

Thus, for the second research question, all three independent variables were related to the amount (percentage of overall errors addressed with oral corrective feedback) and the type (implicit, explicit, or no feedback) of feedback provided in the Spanish FL classroom, although not all differences were significant. NS instructors corrected more and provided explicit feedback significantly more than NNS instructors. Less experienced instructors (those with 6.5 years of teaching experience or less) ignored errors significantly more than instructors with more than 7 years of teaching experience, and those without SLA education corrected significantly more with explicit feedback than those instructors with SLA education.

Research Question 3: Consistency of Feedback Provision

The third RQ investigated potential relationships between the three IDs and the consistency of feedback provision in the SFL classroom. Specifically, it explored the relationship between the type of feedback instructors provided for certain types of errors.
Lexical Errors

Chi-Square comparisons of NS and NNS instructors’ treatment of lexical errors yielded significant results. NSs were significantly more consistent in their provision of explicit feedback for lexical errors when compared to NNSs ($X^2(2) = 23.88, p < .001$). NNSs did not consistently provide one type of feedback more than another when addressing lexical errors, at least not significantly so, $X^2(2) = 2.84, p = .24$.

With regard to SLA education, instructors without SLA education were significantly more consistent with their explicit correction of lexical errors: $X^2(2) = 20.42, p < .001$. Those with SLA education were not consistent, meaning that they did not use one type of feedback significantly more than another: $X^2(2) = 4.90, p = .09$.

With respect to the independent variable of experience, instructors with more than 7 years of experience ($X^2(2) = 13.74, p < .01$) and those with 6.5 years or less ($X^2(2) = 15.99, p < .001$) were consistent: both provided significantly more explicit feedback for lexical errors than implicit and no feedback.

Morphosyntactic Errors

Similar significant result patterns were observed when examining instructors’ treatment of morphological errors for all three independent variables. With respect to NS/NNS status, both NS and NNS instructors were significantly consistent, but in different ways: NS instructors consistently provided explicit feedback for morphological errors ($X^2(2) = 74.58, p < .001$) while NNSs were significantly more likely not to address morphological errors ($X^2(2) = 66.61, p < .001$).

Like NS/NNS status, SLA education yielded significant results in conflicting ways: instructors with SLA education were significantly more likely not to address
morphological errors ($X^2(2) = 34.86, p < .001$) while those without SLA education corrected errors explicitly ($X^2(2) = 46.39, p < .001$).

Finally, there were significant results for instructors of varying years of teaching experience. Those with 6.5 years of experience or less were significantly more likely to not address morphosyntactic errors with corrective feedback ($X^2(2) = 84.66, p < .001$) and those with more than 7 years of experience were at the other extreme: significantly more likely to address the errors with explicit feedback ($X^2(2) = 41.06, p < .001$).

**Phonological Errors**

Unlike instructors’ treatment of lexical and morphosyntactic errors, there were only two instances of significant relationships when examining instructors’ feedback addressing phonological errors: NNS instructors were consistent in the way they did not address phonological errors ($X^2(2) = 27.78, p < .001$), as did instructors with 6.5 years or less of teaching experience ($X^2(2) = 12.83, p < .001$). No other significant relationships were found for instructors who had ($X^2(2) = 3.22, p = .20$) or did not have ($X^2(2) = 1.24, p = .54$) SLA education, for NS instructors ($X^2(2) = 2.06, p = .36$), nor for those instructors with more than 7 years of teaching experience ($X^2(2) = 2.54, p = .28$).

**Incorrect Answers**

Across the board, instructors tended to address students’ incorrect answers, although they differed in the types of feedback they utilized. NS instructors ($X^2(2) = 37.78, p < .001$), those without SLA education ($X^2(2) = 45.08, p < .001$), and those with more than 7 years of teaching experience ($X^2(2) = 31.39, p < .001$) were consistent in their explicit feedback provision for incorrect answers. Instructors who had SLA education were also consistent with their treatment of incorrect answers, although they
used significantly more implicit feedback ($X^2(1) = 6.76, p < .01$). NNS instructors ($X^2(1) = .25, p = .62$) and those with 6.5 years of experience or less ($X^2(1) = .26, p = .61$) were not consistent in addressing incorrect answers.\(^{18}\)

*Unsolicited Use of the L1 (English)*

When it came to addressing unsolicited uses of students’ L1 (English), there were varied results. NNS instructors were consistent in the way they did not address the use of English with feedback ($X^2(2) = 27.12, p < .001$), as were those instructors without SLA education ($X^2(2) = 13.23, p < .001$). Instructors with 6.5 years of teaching experience or less ($X^2(2) = 9.7, p < .01$) were consistent in their provision of explicit feedback for unsolicited uses of the L1. Finally, NS instructors ($X^2(2) = 3.58, p = .17$), those with SLA education ($X^2(2) = 1.19, p = .55$), and those with more than 7 years of teaching experience ($X^2(2) = 2.68, p = .26$) were not consistent in their treatment of L1 use in the Spanish FL classroom, meaning that they did not use a certain type of feedback when dealing with unsolicited student speech in the L1.

*Research Question 3: Summary*

In sum, there were significant relationships between the three individual difference factors and the consistency of feedback provision of certain types of errors. The results can be seen summarized in Table 25 below.

\(^{18}\) Note the degree of freedom (1) for the + SLA instructors, the NNS instructors and those with less experience. This is due to the fact that there is not one instance of “no feedback” for these instructor categories, meaning that the instructors always provided feedback when students gave an incorrect answer.
Table 25. Significant Results in Chi-Square Tests of Feedback Consistency

<table>
<thead>
<tr>
<th></th>
<th>Lexical</th>
<th>Morphosyntactic</th>
<th>Phonological</th>
<th>Incorrect</th>
<th>Use of L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>Explicit ***</td>
<td>Explicit***</td>
<td>Not signif.</td>
<td>Explicit**</td>
<td>Not signif.</td>
</tr>
<tr>
<td>NNS</td>
<td>Not signif.</td>
<td>No fdbk***</td>
<td>No fdbk***</td>
<td>Not signif.</td>
<td>No fdbk***</td>
</tr>
<tr>
<td>- SLA edu.</td>
<td>Explicit ***</td>
<td>Explicit***</td>
<td>Not signif.</td>
<td>Explicit**</td>
<td>No fdbk**</td>
</tr>
<tr>
<td>0-6.5 yrs.</td>
<td>Explicit ***</td>
<td>No fdbk***</td>
<td>No fdbk**</td>
<td>Not signif.</td>
<td>No fdbk**</td>
</tr>
<tr>
<td>&gt; 6.5 yrs.</td>
<td>Explicit **</td>
<td>Explicit***</td>
<td>Not signif.</td>
<td>Explicit**</td>
<td>Not signif.</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
*** p < .001

Research Question 4: Relationship Between Beliefs and Practice

To examine the potential relationship between instructors’ beliefs and practice regarding each topic (i.e., if instructors report that only a few errors should be corrected, how much do they actually correct?), correlations were run. As with the three previous RQs, the independent variables were NS/NNS status, SLA education, and years of teaching experience. For this RQ, the dependent variables were the means of each belief topic and the corresponding practice (see Table 14 from the previous chapter, p. 149). Correlations between each belief topic and behavior revealed the following trends:

Comparing NS and NNS Instructors

For all belief topics and corresponding in-class practices, there were no significant differences for either NS or NNS instructors. The correlations examining each belief topic and corresponding practice are bolded in Table 26; any relationship not bolded
denotes relationships between a non-corresponding belief and practice (e.g., belief about implicit feedback use and actual use of explicit feedback in the classroom), and are thus not of interest to the current study.
Table 26. Comparisons of NS and NNS Instructor Beliefs and Practices

<table>
<thead>
<tr>
<th>Practice</th>
<th>Rate</th>
<th>Implicit</th>
<th>Explicit</th>
<th>Lexical</th>
<th>Morph.</th>
<th>Phon.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>NNS</td>
<td>NS</td>
<td>NNS</td>
<td>NS</td>
<td>NNS</td>
</tr>
<tr>
<td>Amount</td>
<td>-.19</td>
<td>-.15</td>
<td>.17</td>
<td>-.18</td>
<td>-.30</td>
<td>-.14</td>
</tr>
<tr>
<td>Importance</td>
<td>.25</td>
<td>.09</td>
<td>.20</td>
<td>-.29</td>
<td>.13</td>
<td>.11</td>
</tr>
<tr>
<td>Implicit</td>
<td>.15</td>
<td>.18</td>
<td>.35</td>
<td>.44</td>
<td>.03</td>
<td>-.24</td>
</tr>
<tr>
<td>Explicit</td>
<td>.13</td>
<td>-.11</td>
<td>-.32</td>
<td>-.39</td>
<td>.20</td>
<td>.29</td>
</tr>
<tr>
<td>Lexical</td>
<td>-.39</td>
<td>.14</td>
<td>-.30</td>
<td>.44</td>
<td>-.13</td>
<td>-.09</td>
</tr>
<tr>
<td>Morph.</td>
<td>.14</td>
<td>.09</td>
<td>.30</td>
<td>-.35</td>
<td>-1.0</td>
<td>.39</td>
</tr>
<tr>
<td>Phon.</td>
<td>.32</td>
<td>.20</td>
<td>.00</td>
<td>.59**</td>
<td>.29</td>
<td>-.38</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$
Although there were no significant correlations between beliefs and practices, the lack of relationships warrants detailed analysis.

For the belief topic *amount of feedback* and the corresponding practice (% of errors addressed with feedback) correlations of both NS instructors ($r = -0.19$) and NNSs ($r = -0.15$) were not significant. Additionally, both $r$-values were negative, indicating that instructors’ beliefs were inversely related to their in-classroom behavior, albeit very weakly related. Although both NS and NNS instructors believed that feedback provision should be restricted to the linguistic target or to errors that impede communication, their high rates of correction did not correlate with those beliefs; NSs provided corrective feedback for 74.7% of errors (SD = 20), and NNSs did so 50% of the time (SD = 22).\(^{19}\)

Significant correlations were not found for the belief topic and corresponding practice *importance of feedback* according to NS/NNS status either ($r = 0.25$ for NS, $r = 0.09$ for NNS). Although instructors agreed that feedback provision was important, a significant relationship was not found between this belief and rate of error correction.

In correlations comparing beliefs and practices regarding the use of implicit feedback, $r$-values were much higher but still not significant: $r = 0.35$ for NS instructors, 0.44 for NNS. Although both NS and NNS instructors in this study believe they use implicit feedback often (second only to using a combination of explicit and implicit feedback), the data indicate that NS instructors only used implicit feedback for an average of 27.3% (SD = 14.7) of their feedback provision and NNS instructors an average of 26.4% (SD = 17.5). The opposite was true for explicit feedback: while NS and

---

\(^{19}\) Though the errors and feedback in the class recordings were not coded for linguistic target, based on the SR comments, these high rates of correction (74.7% and 50%) included numerous instances where feedback was provided for errors that were not linguistic targets. Thus, although not quantitatively tested in the current study, the qualitative findings corroborate the results from this particular correlation.
NNS instructors both believed they used explicit feedback least often, NNS instructors provided explicit feedback 25.7% (SD 16.4) of the time, NSs used almost double the amount of explicit feedback, 47.8% (SD = 21) of the time. For NSs, the correlation for beliefs and use of explicit feedback was \( r = .20 \), for NNSs, \( r = .29 \).

For the treatment of lexical errors, the correlation \( r \)-value was higher for NSs \( (r = -.48) \) than for NNSs \( (r = .14) \), although the correlation was negative for NSs, citing evidence that their beliefs and correction of lexical errors are inversely related. As reported in RQ1, both NS and NNS instructors believed that lexical errors were less important to correct than morphosyntactic errors but more important than phonological errors. In practice, NS instructors corrected lexical errors an average of 79% of the time (SD = 24.9) while NNSs corrected 62.5% (SD = 26.1). While both NSs and NNSs believed that morphosyntactic errors were the most important to correct, NSs corrected 69.3% of the errors (SD = 28.1) and NNSs 50.3% (SD = 29.1). This lack of relationship is reflected in the low correlations for both NSs \( (r = .16) \) and NNSs \( (r = .22) \) when it came to morphological errors, as well as for the relationship between beliefs and in-class treatment regarding phonological errors; for NSs \( r = .31 \), for NNSs \( r = -.02 \). NS instructors corrected an average of 74.4% of phonological errors (SD = 25.5), NNSs 28.6% (SD = 35.9) although they both claimed they were the least important to correct.

Comparing Instructors With and Without SLA Education

The correlations comparing instructors with and without SLA education reveal one significant relationship in terms of belief and practice. All correlations between corresponding beliefs and practices are bolded in Table 27 below.
### Table 27. Comparing Beliefs and Practices of Instructors With and Without SLA Education

<table>
<thead>
<tr>
<th>Practice</th>
<th>Rate</th>
<th>Implicit</th>
<th>Explicit</th>
<th>Lexical</th>
<th>Morph.</th>
<th>Phon.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ SLA</td>
<td>-SLA</td>
<td>+ SLA</td>
<td>-SLA</td>
<td>+ SLA</td>
<td>-SLA</td>
</tr>
<tr>
<td>Beliefs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount</td>
<td>-.11</td>
<td>-.29</td>
<td>.51*</td>
<td>-.35*</td>
<td>-.51*</td>
<td>-.14</td>
</tr>
<tr>
<td>Importance</td>
<td>.03</td>
<td>.06</td>
<td>.03</td>
<td>-.08</td>
<td>.02</td>
<td>-.02</td>
</tr>
<tr>
<td>Implicit</td>
<td>-.26</td>
<td>.29</td>
<td>.54*</td>
<td>.31</td>
<td>-.39</td>
<td>.04</td>
</tr>
<tr>
<td>Explicit</td>
<td>.28</td>
<td>-.09</td>
<td>-.59*</td>
<td>-.18</td>
<td>.47</td>
<td>.10</td>
</tr>
<tr>
<td>Lexical</td>
<td>-.22</td>
<td>-.25</td>
<td>-.01</td>
<td>.15</td>
<td>-.04</td>
<td>-.32</td>
</tr>
<tr>
<td>Morph.</td>
<td>.25</td>
<td>.06</td>
<td>.12</td>
<td>-.22</td>
<td>.07</td>
<td>.19</td>
</tr>
<tr>
<td>Phon.</td>
<td>.53*</td>
<td>.23</td>
<td>.09</td>
<td>.28</td>
<td>.53</td>
<td>.03</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$
As seen in Table 27, instructors with SLA education had corresponding beliefs and practices regarding implicit feedback that reached levels of significance ($r = .54$; a strong correlation). That is to say, what instructors reported in terms of their use of implicit feedback was in fact what was observed on the classroom recordings: they believed they used implicit feedback often (less than a combination of implicit and explicit feedback but more than explicit feedback alone), and they did, 28.7% of the time (SD = 19.3). This was not the case for explicit feedback ($r = .47$ for + SLA education; $r = .10$ for - SLA), or for any of the other correlations of interest (bolded in Table 27). For explicit feedback, although both + SLA and - SLA instructors believed they used explicit feedback the least, they used it quite frequently (30.1% of the time for + SLA instructors, SD = 22.2, and 44.6% of the time for - SLA instructors, SD = 21.0).

For correlations addressing the rate of feedback provision (i.e., % of errors corrected) and beliefs about amount, $r$-values were negative for instructors with ($r = -.11$) or without ($r = -.29$) SLA education. Both + SLA and - SLA instructors believed feedback should be reserved for errors that impede communication or were the linguistic target for the lesson but they corrected much more often. Instructors with SLA education corrected 58.8% of errors (SD = 26.8) and those without SLA education corrected 70.4% (SD = 21.6). For the correlations for rate of feedback provision and beliefs regarding the importance of feedback, $r$-values were even smaller, .03 (+ SLA education) and .06 (-SLA education). These correlations were positive, however, meaning that the beliefs and practice were in the same direction (instructors agreed oral corrective feedback was important and provided a lot of error correction), although the $r$-value was almost zero.
Similar to results in the NS vs. NNS comparison above, correlations examining the relationship between beliefs and practices regarding the treatment of lexical errors for instructors with and without SLA education yielded negative *r*-values; -.46 for + SLA instructors and -.08 for instructors without SLA education. + SLA instructors corrected lexical errors 63.2% of the time (SD = 31.6) and - SLA instructors corrected 77.8% (SD = 22.6) although both believed lexical errors were not as important to correct as morphosyntactic errors. This negative relationship was not observed for the treatment of morphosyntactic errors (.31 and .12 for + SLA and - SLA instructors, respectively), or for phonological errors (for both groups of instructors, the *r*-value was .24). Instructors with SLA education corrected morphosyntactic errors 54.2% of the time (SD = 32.1) and corrected 53% (SD = 38.1) of phonological errors. Those without SLA corrected morphosyntax 66.8% (SD = 28.1) and phonology 59.1% (SD = 36.4). Thus, although instructors believed that morphosyntactic errors were most important to correct, followed by lexis and phonology, both + SLA and - SLA instructors corrected lexis most often, followed by morphosyntax and then phonology.

*Comparing Instructors With More or Less Teaching Experience*

Unlike the other two independent variables, correlations investigating beliefs and practices of instructors with more (> 6.5)\(^{20}\) or less (0-6.5) years of teaching experience revealed several significant findings (see Table 28).

\(^{20}\) The reader is reminded that for the “more” experienced instructor category, there were no instructors whose experience fell between 6.5 and 7 years of experience.
Table 28. Comparing Beliefs and Practices of Instructors According to Teaching Experience

<table>
<thead>
<tr>
<th>Practice</th>
<th>Rate</th>
<th>Implicit</th>
<th>Explicit</th>
<th>Lexical</th>
<th>Morph.</th>
<th>Phon.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-6.5</td>
<td>&gt; 6.5</td>
<td>0-6.5</td>
<td>&gt; 6.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beliefs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount</td>
<td>-.23</td>
<td>-.06</td>
<td>.00</td>
<td>.04</td>
<td>.26</td>
<td>-.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.31</td>
<td>-.14</td>
<td>-.14</td>
<td>-.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.14</td>
<td>-.09</td>
<td>.07</td>
<td>.04</td>
</tr>
<tr>
<td>Importance</td>
<td>-.09</td>
<td>.40</td>
<td>-.09</td>
<td>.07</td>
<td>-.13</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.16</td>
<td>.07</td>
<td>-.12</td>
<td>.16</td>
</tr>
<tr>
<td>Implicit</td>
<td>.20</td>
<td>-.14</td>
<td>.53**</td>
<td>-.05</td>
<td>.19</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.07</td>
<td>-.20</td>
<td>.19</td>
<td>-.32</td>
</tr>
<tr>
<td>Explicit</td>
<td>.01</td>
<td>.10</td>
<td>-.50**</td>
<td>-.03</td>
<td>.38*</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.08</td>
<td>-.11</td>
<td>.02</td>
<td>-.00</td>
</tr>
<tr>
<td>Lexical</td>
<td>.04</td>
<td>-.68**</td>
<td>.13</td>
<td>-.02</td>
<td>-.09</td>
<td>-.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.03</td>
<td>-.61*</td>
<td>.03</td>
<td>-.52*</td>
</tr>
<tr>
<td>Morph.</td>
<td>.11</td>
<td>.27</td>
<td>-.13</td>
<td>.20</td>
<td>.21</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.24</td>
<td>.06</td>
<td>.16</td>
<td>.33</td>
</tr>
<tr>
<td>Phon.</td>
<td>.39*</td>
<td>.22</td>
<td>.35</td>
<td>-.04</td>
<td>.18</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.12</td>
<td>.35</td>
<td>.33</td>
<td>.14</td>
</tr>
</tbody>
</table>

* $p < .05$
** $p < .01$
*** $p < .001$
For both types of feedback, implicit and explicit, there were significant findings for those less experienced instructors. For use of implicit feedback, correlation results for instructors with 0-6.5 years of experience were strong ($r = .53$), and significant with $p < .001$; the correlation for instructors with more experience was not significant, $r = -.05$.

The same pattern was found for the use of explicit feedback: the correlation for instructors with less teaching experience was moderate and significant, $r = .38$ ($p < .05$) and the correlation for those with more experience was not significant and negative ($r = -.07$). When it came to implicit feedback, instructors, regardless of experience, believed that they used implicit feedback less than a combination of both implicit and explicit feedback but more than explicit feedback alone. But while instructors with less experience used implicit feedback 24.4% (SD = 16.5), more experienced instructors used it 30.6% (SD = 14.3), a small difference but a significant one nonetheless. When it came to the use of explicit feedback, less experienced instructors provided it for 35.4% of errors (SD = 22.3) and more experienced instructors 45.9% (SD = 21.3).

There were also significant results for the treatment of lexical errors, although this time for the more experienced instructors, with an even stronger correlation, $r = -.61$ ($p < .05$); the correlation $r$-value for less experienced instructors was not significant, although it was also negative ($r = -.03$). As with the other two instructor variables, instructors with more or less experience both believed that morphosyntactic errors were the most important to correct, followed by lexical errors and phonological ones. In terms of actual error correction, more experienced instructors corrected 76.8% of lexical errors (SD = 29.6) and less experienced instructors 70.3% (SD = 24.4). For the treatment of both morphological ($r = .16$) and phonological errors ($r = .21$), correlations were not
significant for the less experienced instructors, and likewise for the more experienced instructors ($r = .33$ and .26, respectively). Instructors with more experience corrected morphology $72.7\%$ of the time ($SD = 27$), phonology $75.3\%$ ($SD = 29.4$) and those with less experience addressed morphological errors $55.5\%$ of the time ($SD = 30.1$) and phonological errors $44.8\%$ ($SD = 36.4$). Thus, more experienced instructors corrected lexical errors the most, then phonology, and addressed morphology the least of all the error types, the complete opposite of what they believed.

**Research Question 4: Summary**

To summarize the findings of RQ4, there were very few instances of significant correlations between beliefs and practices, meaning that what the instructors in this study believed and what they did was not consistent. The only significant findings were for + SLA and less experienced instructors for implicit feedback, for less experienced instructors regarding explicit feedback, and for more experienced instructors regarding the error correction of lexical errors.

**Research Question 5: Reasoning Behind Feedback Decisions**

To answer the fifth and final research question, the stimulated recall protocols were examined in detail. The percentage of errors noticed was calculated for each participant as well as the percentage of noticed errors that instructors reported reflecting upon (i.e., noticed errors that instructors consciously thought about correcting or not). These calculations can be seen in Table 29 and are discussed in the following sections.
Table 29. Stimulated Recall Quantitative Data

<table>
<thead>
<tr>
<th>ID</th>
<th>NS</th>
<th>Education</th>
<th>Experience</th>
<th># Errors</th>
<th>Noticed</th>
<th>Conscious</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Yes</td>
<td>6.5 years</td>
<td>10</td>
<td>5 (50%)</td>
<td>2 (40%)</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>11</td>
<td>11 (100%)</td>
<td>10 (91%)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>Yes</td>
<td>3 years</td>
<td>14</td>
<td>14 (100%)</td>
<td>13 (93%)</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>No</td>
<td>5 years</td>
<td>9</td>
<td>9 (100%)</td>
<td>9 (100%)</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>No</td>
<td>6 years</td>
<td>9</td>
<td>6 (67%)</td>
<td>3 (50%)</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>No</td>
<td>30 years</td>
<td>11</td>
<td>8 (73%)</td>
<td>2 (25%)</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>Yes</td>
<td>5 years</td>
<td>8</td>
<td>8 (100%)</td>
<td>8 (100%)</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>No</td>
<td>1 month</td>
<td>9</td>
<td>3 (33%)</td>
<td>3 (100%)</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>10</td>
<td>5 (50%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>10</td>
<td>7 (70%)</td>
<td>7 (100%)</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>Yes</td>
<td>20 years</td>
<td>5</td>
<td>5 (100%)</td>
<td>5 (100%)</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>Yes</td>
<td>1 month</td>
<td>9</td>
<td>4 (44%)</td>
<td>4 (100%)</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>7 (50%)</td>
<td>3 (43%)</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>9 (90%)</td>
<td>5 (56%)</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>No</td>
<td>28 years</td>
<td>10</td>
<td>7 (70%)</td>
<td>1 (14%)</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>No</td>
<td>5 years</td>
<td>11</td>
<td>2 (18%)</td>
<td>1 (50%)</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>5 (50%)</td>
<td>2 (40%)</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>Yes</td>
<td>8 years</td>
<td>11</td>
<td>7 (64%)</td>
<td>5 (71%)</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>6 (60%)</td>
<td>6 (100%)</td>
</tr>
<tr>
<td>34</td>
<td>0</td>
<td>Yes</td>
<td>4 years</td>
<td>10</td>
<td>8 (80%)</td>
<td>3 (38%)</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>No</td>
<td>6 years</td>
<td>10</td>
<td>4 (40%)</td>
<td>3 (75%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>9 (90%)</td>
<td>7 (78%)</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>No</td>
<td>20 years</td>
<td>7</td>
<td>4 (57%)</td>
<td>4 (100%)</td>
</tr>
<tr>
<td>38</td>
<td>1</td>
<td>Yes</td>
<td>10 years</td>
<td>9</td>
<td>1 (11%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>No</td>
<td>1 month</td>
<td>12</td>
<td>7 (58%)</td>
<td>2 (29%)</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>No</td>
<td>2 years</td>
<td>10</td>
<td>2 (20%)</td>
<td>1 (50%)</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>No</td>
<td>8 years</td>
<td>9</td>
<td>7 (78%)</td>
<td>7 (100%)</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>No</td>
<td>3 years</td>
<td>11</td>
<td>7 (64%)</td>
<td>1 (14%)</td>
</tr>
<tr>
<td>46</td>
<td>1</td>
<td>No</td>
<td>4 years</td>
<td>10</td>
<td>1 (10%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
<td>Yes</td>
<td>7 years</td>
<td>11</td>
<td>7 (64%)</td>
<td>3 (43%)</td>
</tr>
<tr>
<td>49</td>
<td>1</td>
<td>Yes</td>
<td>2 years</td>
<td>10</td>
<td>5 (50%)</td>
<td>1 (20%)</td>
</tr>
<tr>
<td>50</td>
<td>1</td>
<td>No</td>
<td>7 years</td>
<td>9</td>
<td>6 (67%)</td>
<td>3 (50%)</td>
</tr>
</tbody>
</table>

After calculating the above-mentioned percentages, regressions were run to see if there were statistical differences in (a) the percentage of noticed errors and (b) the percentage of reflected errors according to each ID factor. Regressions were needed as the dependent variables (the above-mentioned percentages) were continuous, and the independent variables (the three ID factors) were dichotomous.

The researcher then coded the factors that each participant took into account when deciding whether or not to provide feedback. Factors were organized in the following four categories: internal factors (i.e., learner affective factors, prior knowledge, etc.) cited for errors addressed with feedback; internal factors cited for errors not addressed with feedback; external factors (i.e., timing in the class, number of turns already taken,
linguistic target of lesson, course goals, etc.) cited for errors addressed with feedback; and
external factors cited for errors not addressed with feedback.

Comparing NS and NNS Instructors

NS instructors reported noticing 59% of all errors present in the stimulated recall interaction episodes and stated that, for 51% of those noticed errors, they consciously thought about error correction. Percentages were a bit higher for NNS instructors; 67% of errors were reported as noticed, and for 71% of those noticed errors instructors consciously decided whether or not to provide oral corrective feedback. Regressions revealed no significant differences for percentage of noticed errors ($p = .36$) and percentage of noticed errors reflected upon ($p = .14$) for NS/NNS status.

NS Instructors

For the 15 NSs, there were many factors cited as reasons why instructors decided to provide (or not provide) oral corrective feedback (see Table 30). In the Table, the number in parentheses alongside each factor is the number of participants who mentioned that factor, not the number of times cited overall. The bolded factors are the factors mentioned only by NSs and not by NNSs; anything not in bold was mentioned by both.

Table 30. Factors Cited in NS Instructor Stimulated Recalls

<table>
<thead>
<tr>
<th>Internal Factors + Feedback</th>
<th>Internal Factors - Feedback</th>
<th>External Factors + Feedback</th>
<th>External Factors - Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concern for feelings (4)</td>
<td>Lack of ability (4)</td>
<td>Target item (11)</td>
<td>Focus on meaning (7)</td>
</tr>
<tr>
<td>Excellent ability (3)</td>
<td>Concern for feelings (3)</td>
<td>Unexpected answer (6)</td>
<td>Not a ‘serious’ error (1)</td>
</tr>
</tbody>
</table>
The breakdown of factors cited by NS instructors is as follows: NSs cited 7 internal factors that encouraged them to provide feedback, and 6 internal factors that caused them to not provide feedback. There were 8 external factors cited as reasons why feedback was provided and another 8 as justification why errors were not addressed.

As seen in the Table, the most frequent internal factor cited as a reason why instructors provided feedback was their concern for student feelings. Instructors often stated that they provided feedback when learners were struggling with an item in order to stop the student from struggling in front of their peers. The second-most cited internal factor was the opposite: students’ excellent ability with respect to the target language. Many instructors stated that they corrected proficient students because they felt that those particular students would be able to make use of the feedback. Additional internal factors cited as justification for feedback provision included: students’ lack of knowledge of the

<table>
<thead>
<tr>
<th>Lack of knowledge (2)</th>
<th>Developmental readiness (2)</th>
<th>Interrupted meaning (6)</th>
<th>Desire to change stereotype (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared L1 (1)</td>
<td>Lack of understanding of activity, not item (1)</td>
<td>Previous teaching experience (3)</td>
<td>Timing within lesson (1)</td>
</tr>
<tr>
<td>Learning style (1)</td>
<td>Prior language experience (1)</td>
<td>Use of L1 (3)</td>
<td>Instructor knowledge (1)</td>
</tr>
<tr>
<td>Confident student (1)</td>
<td>Student personality (1)</td>
<td>Major error (3)</td>
<td>Test performance (1)</td>
</tr>
<tr>
<td>Student intention (1)</td>
<td></td>
<td>Student pause (2)</td>
<td>Too many mistakes (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Timing within lesson (2)</td>
<td><strong>Teaching method</strong> (1)</td>
</tr>
</tbody>
</table>
target structure (there was no way the student could be expected to correctly produce that particular item without assistance); the instructor and students’ shared knowledge of the L1 (knowing that there was a marked difference from the student’s L1 and thus the instructor wished to address the item); a particular student’s learning style (an instructor stated that she needed to correct that student more often because she did not do as well with aural input); students’ personalities (referring to confident students); and an instructor’s interpretation of what the student was intending to say (comments such as “I knew what they were trying to get across but they just needed a push”). Many of the same internal factors were cited as justification for why not to provide corrective feedback.

The most-cited internal factor for not providing feedback was students’ lack of language ability and developmental readiness (term actually used by one instructor); instructors did not think feedback would have been beneficial for particular student(s) at the time of the error. Instructors also expressed concern for students’ feelings and a student’s shy personality, often citing it as the reason why they did not want to correct certain students. Errors that instructors interpreted being caused by prior foreign language experience (most often French) as well as a lack of understanding of the activity at hand, rather than the linguistic target, were factors that led to a lack of error correction.

NS instructors cited even more external factors as reasons why they did or did not provide oral corrective feedback during class time. Almost all cited that they corrected errors that were the target linguistic item and many cited that they corrected an error that was an unexpected answer or an error that interrupted the meaning of what the student was trying to communicate. Additional external factors that encouraged instructors’ error correction included their previous teaching experience (knowing that a certain item was
more successfully learned with feedback, or that a certain error needed to be corrected prior to continuing on with more complicated structures), student use of the L1 (citing the use as evidence that the student needed assistance), the presence of what an instructor believed to be a ‘serious’ error, timing within the lesson (when there was an opportunity to correct without interrupting the flow of the class) and a student pause, which NS instructors interpreted as a signal of a breakdown in communication as well as a signal that the student was asking for their assistance.

The external factor overwhelmingly cited for not providing feedback was instructors’ focus on meaning at the time of the error. Other factors included: the interpretation that the particular error was not a ‘serious’ error; the instructors’ desire to change a student’s cultural stereotype; the timing of the error within the lesson (one instructor did not correct during the pre-task stage, since students were showing courage by volunteering to try something new); and instructor’s knowledge (one example was an instructor who was unsure how to correct a certain error since the grammar explanation in the student’s text was different than the instructor’s own dialect). One NS instructor also reported that he did not correct a student for a particular error since in the class prior to the one recorded the student produced the form perfectly on an exam. Another did not provide feedback due to the fact that that particular student was making too many mistakes to address any of the errors. Finally, one NS instructor stated that the communicative teaching approach she followed would not encourage correction for a ‘minor’ gender-agreement error.
**NNS Instructors**

The 17 NNS instructors cited even more factors than the NSs (see Table 31). The NNSs cited 6 internal factors that encouraged them to provide error correction and 6 that inhibited them from correcting students. In terms of external factors, NNS instructors cited 15 factors in the + feedback category and 11 in the - feedback category. Any factors in bold were only mentioned by the NNS instructors.

*Table 31. Factors Cited in NNS Instructor Stimulated Recalls*

<table>
<thead>
<tr>
<th>Internal Factors + Feedback</th>
<th>Internal Factors - Feedback</th>
<th>External Factors + Feedback</th>
<th>External Factors - Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ability (5)</td>
<td>Avoid anxiety (5)</td>
<td>Lesson target (5)</td>
<td>Focus on meaning (4)</td>
</tr>
<tr>
<td>Perceived student intention (5)</td>
<td>Student below average (3)</td>
<td>Looked at instructor (5)</td>
<td>Instructor competency/performance (3)</td>
</tr>
<tr>
<td>Prior language experience (3)</td>
<td>Concern for feelings (2)</td>
<td>Interrupted meaning (3)</td>
<td>Waited for student to finish, then forgot (3)</td>
</tr>
<tr>
<td>Frequent error for that student (2)</td>
<td>Personality (2)</td>
<td>Expectations (3)</td>
<td>Not lgx. target of day (2)</td>
</tr>
<tr>
<td>Personality (1)</td>
<td>Frequent error (1)</td>
<td>Looked confused (3)</td>
<td>Sharing personal info. (2)</td>
</tr>
<tr>
<td>Avoid feelings (1)</td>
<td>Adult student (1)</td>
<td>Pause (3)</td>
<td>Not goal of activity (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Questioning tone (2)</td>
<td>Performed structure well on recent exam (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Desire for students to hear correct form (2)</td>
<td>Common error eliminated in time (1)</td>
</tr>
</tbody>
</table>
For the internal factors, NNS instructors cited students’ excellent language ability as the largest reason they provide corrective feedback. Instructors also reported a desire to assist students in their intended communication, and to address errors they believed to be caused by the students’ previous (French) foreign language experience. Additional factors in this category include the student’s confident personality (thus being able to take feedback without it affecting their self-esteem), to help prohibit a student’s feelings from getting hurt by their struggling for the correct form, and the fact that the error was a frequent challenge for that particular student.

Internal factors cited by NNS instructors as explaining why they did not correct particular errors included: a desire to avoid inducing anxiety (the most frequently cited factor); a particular student’s performance being below the class average; concern for a student’s feelings and shy personality; the fact that the student committed that same error
so often, and; the consideration of what correcting a middle-aged student would do in a roomful of young adults.

NNS instructors reported more than twice as many external factors in support of their feedback provision. Instructors stated that they provided feedback for errors that were the linguistic target, when students looked at them, when the error interrupted the communicative meaning of the utterance, when the error was unexpected and when the student produced the error with a questioning tone. Some instructors expressed a desire for students to hear the correct answer and others said they provided feedback when students resorted to using English. NNS instructors also said they corrected a particular error due to the opportune time in a lesson (at the end of one task, prior to beginning the next), because the error had already been discussed in class, or was one of the course objectives (the correct use of the subjunctive for the intermediate level). Instructors also corrected errors perceived as common for the language level, and once an instructor provided feedback because he realized he made a mistake while teaching a structure and the student incorporated the incorrect form.

There were a comparable number of external factors mentioned as reasons why instructors did not provide feedback for particular errors. Most said they were focusing on meaning at that moment and were not concerned with the specificity of the student error. The NNS instructors also expressed a concern about their knowledge and performance and at times waited for a student to finish their thought and then by the time they finished it was too late to correct them in their opinion. If the error was not the linguistic target of the day or the activity, if the student was sharing personal information or had successfully produced the structure on a recent exam, instructors did not correct
the errors. An NNS also reported not correcting a noun’s gender due to the fact he believed the error was common and would eliminate itself over time, while another cited a fear of disrupting the class dynamic. Finally, NNSs described not correcting an error when it came at an inconvenient time during the lesson (during the pre-task explanation), or when the student simply had too many errors in their speech to the point where the instructor did not believe the student would have benefited from feedback.

Factors Unique to NS and NNS Instructors

As demonstrated in the bolded factors in the above Tables (30 and 31), NS and NNS instructors cited many of the same internal and external factors. However, several factors were unique to NS instructors and NNSs, which will be discussed in the following paragraphs. Overall, NS instructors cited 11 factors that NNSs did not mention, and NNS instructors cited 21 factors that were not mentioned by NSs. Like the previous paragraphs, these factors were categorized in terms of internal and external factors, as well as factors that were reported in support of feedback provision and those cited as reasons why feedback was not provided at a given moment.

With respect to internal factors cited in support of feedback provision, NS instructors were the only ones to cite the student’s lack of knowledge or learning style, as well as the fact that both the NS instructor and the student had shared knowledge of the student’s L1 (English). NNS instructors, on the other hand, cited the student’s prior language experience (in French), and the fact that the student made a particular mistake often. When it came to internal factors cited for errors not addressed with feedback, NS instructors cited: a student’s lack of ability; the student’s level of developmental readiness prohibiting correction (which they believed would not be effective); a student’s
lack of understanding of the activity, rather than the item; and the student’s prior language experience (in French). NNSs cited that they wished to avoid producing anxiety for the student, that a particular student was below average and thus the instructor did not want to continuously correct him, as well as the fact that the student made frequent errors with the structure. Finally, a NNS instructor did not correct a student older than the rest of the class for fear of what it would do to that student and to the class dynamic.

With respect to the external factors, there were even more factors unique to the instructors’ NS/NNS status. For external factors cited in support of feedback provision, NS instructors reported taking into account their previous teaching experience (knowing that a certain item needed error correction in order to be learnt), and NSs also reported corrected what they considered to be a ‘major’ error. NNS instructors cited many more external factors that NS instructors did not report. NNSs reported correcting if they received one of three visual or auditory cues: if the student who committed the error looked at them, if the student looked confused, or if the student had a questioning tone in their speech. NNS instructors also corrected when they desired students to hear the correct answer, if it was a topic already covered in class, if the error was something within the course objectives or was a common error. Finally, NNSs corrected if they believe the error interrupted the flow of the classroom or if they had provided the student with incorrect information earlier.

Regarding external factors cited as reasons why instructors did not correct, NS instructors had several factors unique to their NS status, although there were many more reported by NNS instructors only. NS instructors said they did not correct errors they did not consider to be ‘serious,’ when they were more concerned with addressing a cultural
stereotype, or when they believed the error would not be considered important to correct by the communicative teaching approach they followed. NNS instructors, on the other hand, did not correct when they waited too long for a student to finish a thought (causing the instructor to forget that they originally wanted to correct an error), when the error was not a linguistic target of the day or a goal of the activity, when the student was sharing something personal, when the correction would interrupt or negatively impact the class dynamic, or when the error was perceived to be a common mistake eliminated over time.

Comparing Instructors With and Without SLA Education

Instructors with SLA education reported noticing 70% of all errors present in the stimulated recall interaction episodes and reflected upon 66% of the noticed errors. Percentages were a bit lower for those instructors without SLA education: 62% of errors were reported as noticed, and for 65% of those noticed errors instructors consciously decided whether or not to provide oral corrective feedback. These differences were not significant: \( p = .38, p = .79 \), respectively.

Both instructors with and without SLA education reported 8 internal factors explaining their provision of feedback; + SLA instructors cited 7 internal factors as reasons why they did not provide feedback for those particular errors while the - SLA instructors reported 4. Regarding the external factors, + SLA instructors cited 13 factors for feedback provision and 14 for why they did not correct; the external factors for instructors without SLA education were similar in number for the provision of feedback — 14 — but not for explaining those instances why feedback was not provided (only 8 factors were cited by the - SLA instructors).
**Instructors With SLA Education**

Internal factors cited for feedback provision by instructors with SLA education are reported in Table 32. Any factors in bold were unique to the + SLA instructors.

*Table 32. Factors Cited in + SLA Instructor Stimulated Recalls*

<table>
<thead>
<tr>
<th><strong>Internal Factors</strong> + Feedback</th>
<th><strong>Internal Factors</strong> - Feedback</th>
<th><strong>External Factors</strong> + Feedback</th>
<th><strong>External Factors</strong> - Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ability (3)</td>
<td>Avoid anxiety (3)</td>
<td>Linguistic target (9)</td>
<td>Focus on meaning (6)</td>
</tr>
<tr>
<td>Perceived student intention (3)</td>
<td>Perceived student intention (1)</td>
<td>Expectation (4)</td>
<td>Not lgx. target (4)</td>
</tr>
<tr>
<td>Previous language experience (3)</td>
<td>Student feelings (3)</td>
<td>Student look (4)</td>
<td>Teacher performance (3)</td>
</tr>
<tr>
<td>Student knowledge (1)</td>
<td>Student below average (3)</td>
<td>Interrupted meaning (3)</td>
<td>Timing in class (2)</td>
</tr>
<tr>
<td>Confident student (1)</td>
<td>Low confidence (2)</td>
<td>Desire for student to produce correct form (3)</td>
<td>Class dynamic (1)</td>
</tr>
<tr>
<td><strong>Student has frequent errors (1)</strong></td>
<td>Aptitude (2)</td>
<td>Significant error (2)</td>
<td>Desire for student to realize error (1)</td>
</tr>
<tr>
<td>Learning style (1)</td>
<td>Developmental readiness (1)</td>
<td>Student pause (2)</td>
<td>Sharing personal info (1)</td>
</tr>
<tr>
<td>Student feelings (1)</td>
<td></td>
<td>Already covered (2)</td>
<td>Difficulties as NS instructor (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Common mistake (1)</td>
<td>Recent test performance (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Timing in lesson (1)</td>
<td>Too many mistakes (1)</td>
</tr>
</tbody>
</table>
These factors included: excellent student ability; perceived student intention; students’ previous language experience (French); student knowledge; the student’s confident personality; the fact that the student had frequent errors; the student’s learning style; and a desire to avoid the student feeling any embarrassment that his classmates were not understanding him. The internal factors cited for errors that were not corrected included: a desire to avoid causing anxiety; the perceived student intention; student feelings; a student’s general performance being below average; a student’s low confidence; and the lack of developmental readiness for that particular structure.

+ SLA instructors cited almost twice as many external factors for their feedback provision as internal. Factors included: the fact that an error was a linguistic target; was unexpected; was an item previously discussed or covered in class or was a common mistake; was an error that interrupted the communicative meaning of the utterance; and that the instructor thought it was a significant error or had a desire for the student to produce the correct form. Instructors also used visual cues to determine which errors to correct, including a student’s look (at the instructor), a pause in speech, or an impression that the student was struggling. Finally, instructors corrected errors if a student had already spoken in class (thus lessening the instructor’s fear that correction would inhibit their speech), when the timing allowed it, and if the error was one of the main course
objectives. Almost the same number of external factors was cited as explaining why instructors did not provide oral corrective feedback for errors from their recorded class.

Once again, the most cited factor was the instructor’s focus on meaning, followed by the fact that an error was not a linguistic target. Instructors also cited a preoccupation with their own performance, that the timing prohibited error correction at that moment, as well as a fear of interrupting the class dynamic. Additional factors included the inopportune placement of the error within the learner’s speech, a desire for the student to realize their own mistake, the fact that the learner was sharing something personal, and a student’s ability to correctly produce the target structure on a recent assessment. Difficulties caused by NS status and lack of understanding, the fact that a learner had numerous mistakes, and the instructor’s desire to address a stereotype rather than an error were also named in support of not providing feedback. Finally, instructors also cited their teaching approach and their previous teaching experience, stating that the communicative approach frowns on correcting minor errors and that, based on their previous teaching experience teaching at this level, the error will work itself out within a couple of weeks.

Instructors Without SLA Education

Instructors who did not have any SLA education reported the following internal factors for errors which they corrected: student ability; student feelings; perception of student understanding; student knowledge; personality; previous language experience; perceived student intention and shared knowledge of the L1. Half of these internal factors were cited as reasons why instructors did not correct errors: avoid student feelings; student ability; previous language experience and student personality (see Table 33).
Table 33. Factors Cited in - SLA Instructor Stimulated Recalls

<table>
<thead>
<tr>
<th>Internal Factors</th>
<th>Internal Factors</th>
<th>External Factors</th>
<th>External Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Feedback</td>
<td>- Feedback</td>
<td>+ Feedback</td>
<td>- Feedback</td>
</tr>
<tr>
<td>Student ability (6)</td>
<td>Avoid feelings (2)</td>
<td>Target item (12)</td>
<td>Focus on meaning (6)</td>
</tr>
<tr>
<td>Student feelings (4)</td>
<td>Student ability (2)</td>
<td>Interrupted meaning (6)</td>
<td>Timing of error (2)</td>
</tr>
<tr>
<td>Perception of</td>
<td>Previous language</td>
<td>Expectations (6)</td>
<td></td>
</tr>
<tr>
<td>student</td>
<td>experience (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>understanding (2)</td>
<td>Student knowledge (2)</td>
<td>Already covered (3)</td>
<td></td>
</tr>
<tr>
<td>Personality (2)</td>
<td>Personality (1)</td>
<td>Use of L1 (3)</td>
<td></td>
</tr>
<tr>
<td>Shared L1 (1)</td>
<td>Previous language</td>
<td>Student pause (3)</td>
<td>Goal of activity (1)</td>
</tr>
<tr>
<td></td>
<td>experience (1)</td>
<td>Looked at instructor (3)</td>
<td>Teaching method (1)</td>
</tr>
<tr>
<td>Perceived student</td>
<td>Perceived student</td>
<td>Appeared to be struggling (2)</td>
<td>Frequent error for that</td>
</tr>
<tr>
<td>intention (1)</td>
<td>intention (1)</td>
<td></td>
<td>student (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Previously teaching</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>experience (1)</td>
</tr>
</tbody>
</table>
These same instructors cited numerous external factors to explain why they provided feedback for certain errors. These factors include the error being the target item, the fact that the error interrupted the meaning or was unexpected, the error had already been covered, the student resorted to the L1, the student paused, looked at the instructor or appeared to be struggling. The instructors also cited their previous teaching experience and knowing that they had to provide corrective feedback, the timing of the activity (pause, transition, would not interrupt the flow of the class), the fact that an error caused an unintentionally disrespectful statement, the ‘seriousness’ of the error, the number of times the error was said, and the tone of the student. External factors cited as reasons why errors were not corrected included: focus on meaning; timing of error; waited for student to finish; listening for linguistic target; previously teaching experience; goal of activity; teaching method; and the fact that the error was a very frequent error for that student (the instructor said “what’s the point!”).

Factors Unique to + SLA and - SLA Instructors

Overall, the + SLA instructors provided more unique factors (19) than those without SLA education (11). The two internal factors cited for reasons why + SLA instructors provided feedback were the student’s learning style and the fact that a particular student has frequent errors of that type. - SLA instructors cited a shared L1 with the student and their perception of the student understanding as reasons why they corrected (for the latter, the instructor felt that the student did not understand what they were actually saying). In terms of internal factors cited as reasons why instructors did not correct, the + SLA instructors cited a desire to avoid anxiety, the fact that they understood what the student was trying to say, and the fact that the particular student was
below average or was not developmentally ready (term used by one of the + SLA instructors). The instructors without SLA education, on the other hand, cited only the learner’s previous language experience (in French).

As with the other independent variable discussed thus far, both instructors with and without SLA education provided many more external factors in explanation of their feedback provision decisions. + SLA instructors cited four external factors in support of their feedback provision: a desire for the student to hear the correct form; the fact that the error was a common mistake; and that the error was important to course objectives. The + SLA instructors also cited the fact that, as a student had already spoken in class, they were not concerned about inhibiting the student’s participation. - SLA instructors detailed five external factors not mentioned by those with education: the student used the shared L1; said something that was disrespectful in the L2; and the fact that the student had already said the error several times or had a questioning tone. One - SLA instructor also stated that one error was corrected based on her previous experience teaching the course.

Regarding external factors cited as reasons why instructors did not correct, those with SLA education reported three times as many factors as those without SLA education. + SLA instructors cited: a concern for their own performance; a desire to change cultural stereotypes held by a student; the fact that there were numerous mistakes in the student’s speech or the student was sharing personal information; the fact that the student produced the item correctly on a recent test; a desire to not interrupt the dynamic; the awkward placement of the error (in the middle of the student’s paragraph-length discourse); a desire for the student to realize her own mistake; and a failure to correct due to reported difficulties as a NS (one instructor reported it was difficult for her to
distinguish between student vowels at a certain developmental stage due to her NS status). SLA instructors cited only three external factors as reasons why they did not correct: they were waiting for a student to finish a thought; the error was not the target of the activity; and the fact that an error was a frequent mistake for a particular student.

**Comparing More and Less Experienced Instructors**

The largest differences were found when instructor reports were examined according to teaching experience. More experienced instructors reported noticing 65% of errors and reflecting upon 56% of them. Less experienced instructors noticed 61% of errors and consciously thought about correction for 63% of the errors. Again, these differences were not significant: $p = .61$, $p = .97$. Although these percentages between the two groups are similar, the factors cited and the comments paint an entirely different picture of the error correction processes of instructors with more or less years of teaching experience (see Tables 34 and 35).

**Table 34. Factors Cited in + Experience Instructor Stimulated Recalls**

<table>
<thead>
<tr>
<th>Internal Factors + Feedback</th>
<th>Internal Factors - Feedback</th>
<th>External Factors + Feedback</th>
<th>External Factors - Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ability (1)</td>
<td>Aptitude (1)</td>
<td>Linguistic target (7)</td>
<td>Focus on meaning (3)</td>
</tr>
<tr>
<td>Perceived student understanding (1)</td>
<td>Perceived student understanding (1)</td>
<td>Interrupted meaning (4)</td>
<td>Change cultural stereotype (1)</td>
</tr>
<tr>
<td>Student feelings (1)</td>
<td>Student intention (1)</td>
<td>Expectation (4)</td>
<td></td>
</tr>
<tr>
<td>Previous language (1)</td>
<td></td>
<td>Use of L1 (2)</td>
<td></td>
</tr>
<tr>
<td>Learning style (1)</td>
<td></td>
<td>Previous teaching experience (2)</td>
<td></td>
</tr>
</tbody>
</table>
### Table 35. Factors Cited in - Experience Instructor Stimulated Recalls

<table>
<thead>
<tr>
<th>Internal Factors + Feedback</th>
<th>Internal Factors - Feedback</th>
<th>External Factors + Feedback</th>
<th>External Factors - Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student ability (9)</td>
<td><strong>Student feelings</strong> (5)</td>
<td>Lg. target (13)</td>
<td>Focus on meaning (12)</td>
</tr>
<tr>
<td>Perceived student intent (5)</td>
<td>Student ability (5)</td>
<td>Expectations (8)</td>
<td>Timing in lesson (3)</td>
</tr>
<tr>
<td>Previous language experience (3)</td>
<td>Avoid anxiety (4)</td>
<td>Interrupted meaning (7)</td>
<td>Lesson focus (2)</td>
</tr>
<tr>
<td>Student feelings (3)</td>
<td><strong>Student personality</strong> (3)</td>
<td><strong>Student look</strong> (6)</td>
<td>Waiting for the right moment (2)</td>
</tr>
<tr>
<td>Student knowledge (2)</td>
<td>Adult student (1)</td>
<td>Common error (6)</td>
<td>Teacher performance (2)</td>
</tr>
<tr>
<td>Anxiety (2)</td>
<td>Previous language experience (1)</td>
<td>Student paused (4)</td>
<td>Recent test performance (2)</td>
</tr>
<tr>
<td>Confident student (2)</td>
<td>Student makes frequent errors (1)</td>
<td>Error timing (3)</td>
<td>Class dynamic (1)</td>
</tr>
<tr>
<td>Shared L1 (1)</td>
<td></td>
<td>Already spoke (2)</td>
<td>Error placement (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Desire for students to hear correct form (2)</td>
<td>Desire for student to work it out (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructor experience (2)</td>
<td>Course objective (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teaching method (1)</td>
</tr>
</tbody>
</table>
More Experienced Instructors

Internal factors mentioned by more experienced instructors for errors that were corrected included: a student’s excellent ability; perceived student intention; a desire to avoid a negative impact on student feelings; a student’s previous language experience and learning style. Internal factors for errors that were not corrected were very limited: a student’s low aptitude; perceived student intention; and a student’s lack of understanding of the activity, rather than an issue with the non-target utterance.

Instructors with more than seven years of experience cited more external factors for corrected errors, including: the error was a linguistic target; the error interrupted the communicative meaning of the utterance; the error was unexpected; and the fact that the student used their L1. Additional external factors reported in support of feedback provision included an instructor’s previous teaching experience, the seriousness of error, the fact that an error had been previously discussed, and a desire for students to work it out. There were only two external factors for errors that were not corrected for the more experienced instructors: a focus on meaning and a desire to change a cultural stereotype.
Less Experienced Instructors

Less experienced instructors (see Table 35) cited almost twice as many factors for their feedback decisions as the more experienced instructors. Internal factors for corrected errors were the following: a student’s proficient ability; perceived student intent; a student’s previous language experience; student feelings; student knowledge; anxiety; a student’s confident personality; and factors relating to an student’s L1. Internal factors for uncorrected errors were: student feelings; a student’s low ability; a desire to avoid an anxiety-causing situation; a student’s shy personality; the fact that student who produced the error was much older than the others; a student’s previous language experience; and the fact that a student made frequent errors.

Less experienced instructors named many external factors for corrected errors: the fact that the error was a linguistic target; was unexpected; interrupted the meaning; was a common mistake; or came at a convenient time during the lesson were causes for correction. Physical or auditory cues such as student look, a pause, or a student’s questioning tone also led to error correction. Instructors also corrected errors when they had a desire for students to hear the correct answer, when their experience dictated correction of that item was necessary, when it was a major course objective or serious error and when the student used the L1. External factors for uncorrected errors included: a focus on meaning; an inconvenient time in the lesson; the error not being a linguistic item; a course objective or the goal of an activity; or if the instructor was waiting for the right moment and the opportunity for correction passed. Instructors also did not correct due to a student having produced the item well on a recent exam, a desire to keep the
class dynamic pleasant, a desire for the student to work it out themselves, or because it was seen as prohibited by the instructor’s teaching method.

Factors Unique to More and Less Experienced Instructors

When examining instructor factors cited in support of feedback decisions, the most differences can be seen when examining instructors according to experience. Overall, more experienced instructors cited 6 factors that the less experienced instructors did not mention, while the less experienced instructors reported 33 factors that the more experienced instructors did not mention.

More experienced instructors cited only 2 internal factors in support of their feedback provision: perceived student understanding and a student’s learning style. Less experienced instructors cited 5 factors in this category: perceived student intent; student knowledge; a student’s anxiety level or personality; and the instructor and students’ shared L1. With respect to internal factors cited as reasons why an error was not addressed with feedback, more experienced instructors cited 2 factors: perceived student understanding and intention. Less experienced instructors cited 6 factors: a concern for student feelings/anxiety level; a student’s shy personality or previous language experience; and the fact that a student made frequent errors or was an older adult.

There are even greater differences when more and less experienced instructors are compared in terms of the unique external factors cited. More experienced instructors again cited only 1 factor: the fact that a certain error was already discussed. Less experienced instructors cited 9 external factors in explanation of why they provided feedback: a student look, tone or pause; the fact that the mistake was common or part of the course objective; the timing of an error was convenient for correction; a desire for
students to hear the correct answer; and the fact that a student had already spoken. A teacher’s concern for their own competency also led them to correct an error, as they did not want the students to be confused due to an earlier explanation that the instructor later decided was inadequate. More experienced instructors also cited only 1 external factor that less experienced instructors did not; those with more experience cited the desire to change a student’s cultural stereotype. Less experienced instructors, on the other hand, cited 13 unique external factors as reasons why they did not correct. Those with less experience did not correct when the timing of the lesson did not permit, when it was not the “right” moment for them to correct, or when the error was in the middle of a student’s speech. Less experienced instructors also did not correct when the error was not the focus of a lesson, goal of the activity, or was not part of the course objective. The fact that a student produced the item correctly on a recent test was cited as a reason why they did not correct an error, as was an instructor’s desire for a student to work out the error on their own. Those with less experience also reported not correcting an error believed to eliminate itself over time, when an error was part of a student’s personal story, or when the error was not considered important to communication (according to the instructor’s communicative teaching approach). Finally, less experienced instructors cited their own performance as reason why, at times, they did not correct students.

Research Question 5: Summary

To summarize the results revealed for the fifth research question, instructors provided many internal and external factors in explanation of why they did or did not correct their students during class time. All reported very high levels of noticing for the errors, although there were differences according to individual variables. Instructors with
SLA education reported noticing the most errors, 70%, and NNSs instructors reported reflecting on the most errors (71%). All instructors reported more external than internal factors, and the fewest factors reported overall were by the more experienced instructors.
CHAPTER FIVE
DISCUSSION AND CONCLUSION

In the following sections, the results from the current dissertation are discussed in tandem with findings from previous interactional feedback and teacher research outlined in the literature review (whenever possible). The reader is reminded that, due to the paucity of research on the topics at hand, null hypotheses were tested for each research question. For ease of presentation, each research question will be restated. After the discussion, the conclusion will summarize the key findings of the study and articulate the theoretical, methodological and pedagogical implications that research into instructor ID factors can have for the field of SLA.

Discussion

Research Question 1: Beliefs

RQ1: Do instructors of foreign language (FL) Spanish differ in their beliefs about the provision of oral corrective feedback in the classroom? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

As reported in the results section and shown in Tables 16-23, linear regressions found no significant relationships between any of the three ID factors – NS/NNS status, SLA education or years of teaching experience – and the instructor belief topics, which are discussed individually in the following paragraphs. Thus, the null hypothesis was retained: instructors of FL Spanish do not differ in their beliefs about oral corrective feedback provision in the classroom. The reader is reminded that each belief topic consists of answers to several questions from the Belief Questionnaire; see Table 7.
Amount of Feedback

All instructors, regardless of ID comparison, believed that feedback should be reserved for the linguistic target of the lesson, or for errors that impeded communication; very few instructors stated that they correct every error or occasional errors. And most disagreed or strongly disagreed with the statement in Q21A, “Instructors have the responsibility to correct all mistakes uttered by students.” Only one instructor agreed with the statement, one neither agreed nor disagreed and nobody strongly agreed with it. The open-ended question associated with this belief topic corroborates the data from the closed questions. Instructors, regardless of ID factor, explained their amount of feedback provision as focusing on errors that are the linguistic target(s) of the lesson or errors that impede communication.

The belief that feedback should be reserved for errors that interfere with communication is something found in other instructor-focused studies such as Schulz (1996, 2001) and Basturkmen et al. (2004); the latter was the only study specifically focusing on instructor beliefs about error correction while the former set focused on aspects of grammar teaching in general. The three instructors in Basturkmen et al. (2004) also stated that correction should be reserved for errors that interfere with the communicative message. Schulz (1996) found that only 42% of instructors felt that spoken errors should be corrected, while more than double (92%) believed that written errors should be addressed.

Importance of Feedback

The second belief topic, investigating instructors’ beliefs about the importance of feedback similarly found no significant differences according to instructor ID: the
overwhelming majority of instructors believed that oral corrective feedback in the classroom is important. Although the overall means for instructors were similar, an in-depth analysis of the specific questions that comprised this belief topic revealed inconsistencies in instructor opinions. Most instructors (83%) reported that they agreed or strongly agreed with the statement “Oral corrective feedback in the classroom is important.” However, when it came to uncovering why instructors believe feedback is important, results were less clear: 48% of instructors agree or strongly agree that without oral corrective feedback students will continue to make the same mistakes; 29% neither agree nor disagree and 23% disagree. And while instructors believe they have a responsibility to provide feedback, they do not necessarily see it as their most important role in the classroom.

For one question (Q21B), instructors were asked to rate their agreement with the statement “One of my major roles as an instructor is to provide oral corrective feedback.” Over 64% of instructors either agreed or strongly agreed with the statement; 16% of instructors neither agreed nor disagreed with the statement, and approximately 20% of instructors disagreed with it. However on another question (Q1B), instructors ranked feedback provision as the fourth (out of five) most important role as a language instructor; motivate students about learning the foreign language was ranked first, followed by providing input in the foreign language and then providing students with cultural information. Only “act as a reference” was ranked lower than feedback provision.

The answers for the open question (Q25) “What is your overall opinion of the use of oral corrective feedback in the classroom?” echo these opinions. While most (58%) instructors who responded said they view oral corrective feedback as “important,” or
“necessary,” all but a couple qualified their statements. Most frequently, instructors expressed concern of using feedback “too much” or “making students feel bad.” More than half of the instructors voiced concerns that they are unsure about the validity of oral corrective feedback and/or they are unsure about how to best provide it.

This dissociation between the importance of feedback and instructors’ role in providing it is perhaps indicative of the fact that feedback is considered, researched, and taught separately from instructor considerations, focusing largely on the learner and the learning environment. And, as discussed in the Instructor Education section of the literature, feedback is usually one of many topics covered in teacher education courses, seldom discussed in great detail, and almost always discussed in terms of learner IDs. It was very clear from the opinion questions (and in the SR protocols, discussed in RQ5) that almost all instructors had a preoccupation with making the students nervous as a result of feedback provision.

These inconsistencies regarding the importance of feedback is another trend found in the literature. In Basturkmen et al. (2004), one instructor stated that he was “very suspicious” of error correction (p. 259) but that he also believed it was important to correct target structures. The other two stated that they found the flow of communication to be paramount and were reluctant to correct for that reason, but that they also believed it was important to stop the class if an error was interrupting communication. Thus, Basturkmen et al.’s findings are congruous with the findings explained here.

Use of Implicit and Explicit Feedback

With respect to instructor beliefs regarding their use of implicit feedback, instructors in this study believed that their use of implicit feedback fell between their use
of explicit feedback (the least frequent type of feedback instructors believe they use) and a combination of explicit and implicit feedback (Q13B). In the open-ended question (Q14), instructors cited implicit feedback’s ability to “seem more natural,” “be more subtle” and “give students a better chance to find the right answer.” Again instructors reported a concern for embarrassing the student or causing anxiety, stating that “I try not to embarrass the student too much,” in support of their use of implicit feedback.

Instructors saw explicit feedback as a much more straight-forward and guaranteed way to capture student attention often used if implicit feedback does not work; one instructor commented, “I try implicitly but …when I am quite sure the student does not understand at all why he/she made a mistake, I go with the explicit explanation.” A couple of instructors reported that they favored explicit feedback for its direct nature, and two stated their preference for it because it is “quicker.”

The vast majority of instructors, no matter their preference for feedback type, qualified their statements with many comments such as “it depends on the situation/the activity/the student/ the setting/ the topic/ the time constraint,” etc. Six did not provide any qualifying statements with respect to the type of feedback they provide, stating simply that they use one type of feedback over another “because I think [implicit] is the best way,” or because “[explicit feedback provision] is just what I do.” Interestingly, all instructors that did not qualify their statement or provide additional factors that affected the type of feedback they provide were NS instructors; some with SLA education, some without, and mixed between instructors with more and less teaching experience.

This less-than-straightforward way of describing what type of feedback should be used and why has been found in other studies, although normally between instructors. In
Basturkmen et al. (2004) for example, two of the three instructors were resolute in their beliefs about the type of feedback to use, though in contrasting ways. One believed that explicit feedback was the best type to use; he did not trust implicit techniques and found recasts too vague to be helpful to the students. The other was very confident that implicit feedback (specifically recasts) was the most effective. The third instructor reported inconsistencies within his own belief system. While he favored student correction and felt that implicit feedback was the preferred type of feedback, when it came to describing how he would alert learners to an error in hopes that the student would then self-correct, the instructor cited very explicit techniques such as a thumbs-down sign or writing on the board; the opposite of implicit feedback. At the same time, all three instructors believed that student-initiated correction was preferable, and that correction should only take place when the error interrupted the learners’ communicative message.

Treatment of Specific Errors: Lexical, Morphosyntactic or Phonological

Instructor beliefs were similar regarding the treatment of specific errors. In this study, instructors, regardless of NS/NNS status, SLA education or teaching experience, believed that morphosyntactic errors are the most important to correct, followed by lexical and phonological errors. When asked if they believe they differ in the way they correct the different error types (Q23), instructors were more or less evenly divided: 58% said they do differ in the way they correct each error type, while 42% said they correct the same across the board. In the open-ended question (Q24), the majority of instructors said they would correct phonological errors by repeating the correct form, either in the form of a recast or repetition or with a more explicit, overt correction; the responses were very evenly split in terms of explicit or implicit use. Two instructors (one NNS without
SLA education and more than 35 years of experience and one NS without SLA education and more than 15 years of experience) stated that they would make the entire class repeat the correct form. For the morphosyntactic question, instructors again cited both implicit and explicit techniques. This time, however, instructors cited several factors that would encourage them to use explicit feedback, including: a frequent error; a linguistic target; or an error that prohibits the communicative meaning of what the student was saying.

Finally, for lexical errors, the majority of instructors said they would provide a recast to provide the correct word, with a third of the instructors stating that if the word had not been studied before they would write it on the board. A couple of the instructors said they would wait until the end of the activity or class period to write all errors, including lexis, on the board so that all students could benefit from the feedback.

The majority of the instructor belief studies discussed in detail in the literature review (e.g., Árva & Medgyes, 2000; Chavez, 2006; Schulz, 1996, 2001), did not ask instructors about their beliefs regarding correction of certain types of errors. The three instructors in Basturkmen et al. (2004) had very different beliefs: one instructor was very focused on correcting vocabulary and phonology, another believed grammar was the most important to correct, while the third thought forms discussed in the previous lesson were the most important to focus on. Thus, the current study differs from Basturkmen et al. in regards to beliefs about error type: instructor beliefs in the current study were very homogeneous regarding the type of errors that should be corrected: morphosyntax, followed by lexis and then phonology.
Research Question 1: Summary

The fact that the 60 instructors in this study, regardless of comparisons of three ID factors, did not significantly differ on any of the seven belief topics is surprising. Studies investigating beliefs have usually found at least some differences between instructors (e.g., Basturkmen et al., 2004; Chavez, 2006), however, both of aforementioned these studies were case studies, making comparisons between results difficult. As discussed in chapter 2, research examining instructor beliefs has traditionally focused on common themes that emerge from the data, the differences between pre- and in-service instructors, or the changes that result over time or after participation in a teacher training course. This lack of research into potential differences between instructors’ beliefs, particularly while considering instructor ID factors, prohibits comparisons with other studies.

From the detailed analyses above it may be the case that instructors in this study did not differ significantly in their beliefs according to ID factor because their beliefs were so inconsistent. Perhaps no significant differences were found because the majority of the instructors, regardless of the ID factor considered, were similarly inconsistent and seemingly contradictory in their beliefs about feedback. The other possibility is, of course, that instructors in this study genuinely did have the same beliefs about oral corrective feedback. This possibility is supported by (a) the low SD values present in the analyses, meaning that the instructor beliefs were similar, or at least similarly inconsistent (as seen in Table 16), and (b) the fact that more than half of the instructors (32) took the same Spanish Teaching Methodology course offered at Georgetown University; 92% of all instructors said they had participated in a formal teaching methods course or teacher training seminar at some point in their career.
Research Question 2: Feedback Provision

RQ2: Do instructors of FL Spanish differ in the type and amount of oral corrective feedback they provide in the classroom? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

Results from the second RQ revealed significant findings for both amount and type of feedback provided. The null hypotheses are therefore rejected: instructors do differ in the type and amount of feedback they provide in the Spanish FL classroom. The amount of feedback provided was related to instructor NS/NNS status, and the type of feedback was related to each of the instructor ID factors.

Amount of Feedback

A linear regression found only instructor NS/NNS status to be related to the percentage of errors addressed with oral corrective feedback (see Table 24). On average, NS instructors corrected 74.7% of all errors, compared to 50% of NNSs. This difference of 24.7% was significant ($p < .01$). There were also differences according to the amount of feedback provided in relation to education (an average difference of 11.6%; $p = .16$), and in relation to experience (an average difference of 16.1%; $p = .08$), although these differences were not significant. With respect to SLA education, instructors without SLA education corrected more than those with SLA education (70.4% compared to 58.8%, respectively); those with more teaching experience corrected more than instructors with less teaching experience (76% and 59.9%). See Figure 2 for these differences (p. 160).

Instructor NS/NNS Status

Although the null hypotheses were tested, the finding that NS instructors corrected significantly more than NNSs is not unexpected if the results from two of the
pilot studies are considered. In the Qualifying Paper study (spring 2007), the four NS instructors corrected an average of 71% of errors, whereas the three NNSs corrected 53% of the time. In the Research Methods study, the NS corrected 85% of the time compared to 60% correction by the NNS.\textsuperscript{21} However, the results from this study were opposite of the findings from instructor error evaluation studies (James, 1977; Sheorey, 1986; Porte, 1999) and comparisons of NS and NNS instructor in-class feedback provision (Árva & Medgyes, 2000) discussed in the literature review; in all of the aforementioned studies NNS instructors corrected more than the NSs. The finding that, in this study, NNSs corrected significantly less than NSs was also the opposite finding expected according to studies investigating student perceptions of how NSs and NNSs differ from each other (e.g., Barratt & Kontra, 2000; Benke & Medgyes, 2005; Lasagabaster & Sierra, 2005); in all of these studies NSs were expected to be more tolerant of errors than NNSs.

Had just the classroom data been collected, it would have been difficult to explain these differences that were opposite of what the literature would predict. Fortunately, the stimulated recall protocol data can explain these differences. NNS instructors reported that they took into account almost twice as many factors as the NSs when making in-class feedback decisions, providing evidence of why they corrected only 50% of errors compared to NS instructors’ 74.7%. This, along with results for all of the observed behaviors, will be discussed in detail with the RQ5 results.

\textit{Instructor SLA Education}

Despite the fact that there were no significant differences according to instructor education or experience, the low \( p \) values warrant discussion. Though there was no

\textsuperscript{21} Results from the third and final pilot are not discussed, as the classes observed in the Pilot were movie discussions and thus incomparable to the others.
empirical research to predict how instructor SLA education would relate to their feedback provision, the finding that instructors with SLA education corrected less than instructors without it makes sense. As will be discussed in the SR comment analysis (RQ5), instructors with SLA education would presumably take many more factors into consideration in relation to their feedback provision. In SLA courses, one is taught about the mediating effects that internal and external factors have on L2 learning in general, as well as on specific aspects such as feedback. It is reasonable to think that + SLA instructors would take some of those factors into account during class time, resulting in less feedback provision when compared to a - SLA instructor who is presumably not thinking about those factors. Additionally, as SLA courses discuss contextual factors that can raise learners’ affective filters, and decrease motivation and willingness to communicate, it seems logical to think that instructors with SLA education could be more concerned about negatively affecting the SLA process than those without SLA education. Although this is the first study to investigate potential differences between instructors who have taken two or more SLA courses and those who have taken none, studies have found that even a one-semester teacher training course affects instructor behaviors (see reviews in Borg, 2003, 2006a) and thus one could expect that having participated in multiple SLA courses would potentially influence instructors’ teaching behaviors.

Instructor Experience

The fact that instructors with more teaching experience corrected much more than those with less experience ($p = .08$) is another logical finding that is supported in part by what the literature would expect. As instructors gain more experience, many of their in-class behaviors become more automatic and there is a decrease in the amount of
reflective in-class decision-making (Berliner, 1995); one would expect this to apply to feedback decision-making as well. The fewer factors taken into account while teaching, the more resources instructors have to act in class, such as provide feedback. Results from Mackey et al. (2004) were in line with those found in this study, that experienced instructors corrected more than the novice instructors. Other studies, however, such as Polio et al. (2006) did not find significant differences in the amount of recasts provided by experienced and pre-service instructors. However, in both of these studies, experienced instructors (with years of experience ranging between 4-15 and 4-27 years, respectively) were compared with instructors who had never taught before, and both studies took place in the ESL context, so comparisons must be made with caution.

Type of Feedback

Chi-square tests of independence discovered that each of the three instructor ID factors was related to their in-class feedback provision. NS instructors gave explicit feedback significantly more than NNSs, instructors with no SLA education corrected with significantly more explicit feedback compared to those with SLA education, and instructors with less teaching experience ignored errors significantly more than those with more teaching experience.

Instructor NS/NNS Status

First, let us examine the significant findings for instructor NS/NNS status. When NS instructors provided feedback, 47.4% of the time it was explicit feedback (27.3% of the time it was implicit). NNS instructors, on the other hand, only provided explicit feedback 23.6% of the time (relying on implicit feedback 26.4%). This was extremely interesting, as other classroom studies have found implicit feedback to be used more
often than explicit (e.g., Sheen, 2004; Lyster & Mori, 2006). However, in these aforementioned studies, details on the instructor were minimal. Sheen (2004) presented the range of the years of teaching experience in each of the four contexts, as well as the variation of languages, but did not provide concrete details. The instructors in Lyster and Mori’s (2006) study were of similarly mixed backgrounds, although the study was detailed in reporting the background attributable to each instructor. Thus, the current study demonstrates the importance of considering each instructor’s ID factors in relation to feedback provision and why lumping all instructors together in analyses could conceal part of the interpretation.

**Instructor SLA Education**

With respect to the ID factor of SLA education, those instructors without it corrected significantly more with explicit feedback than those instructors who had SLA education. When instructors without SLA education corrected errors, 44.6% of the time it was with explicit feedback (compared to 26% implicit). Those with SLA education also used explicit feedback more (30.1%), although only slightly more than implicit feedback (28.7%), and not at the level of significance. As there are no studies investigating instructor SLA education in relation to feedback provision, these results cannot be interpreted in light of other studies. Relying solely on the data from the current study, one could presume that something in the recorded interactions led instructors who have SLA education to correct with more explicit feedback. Perhaps it was something learned in their respective SLA coursework, perhaps not; it is impossible to explain why with only the recorded data. As discussed with RQ1, the belief questionnaires shed no light on any
observed differences. Once again, triangulation proved to be a sound methodological decision and these behavioral differences are discussed further in RQ5.

There is one more consideration that could be addressed in future studies. For the type of feedback, both NS instructors and those without SLA education provided significantly more feedback than their NNS and + SLA counterparts. Although each ID factor was considered separately in the current study thus maintaining the integrity of the statistical analyses, the fact remains that that 29 of the 40 - SLA instructors were NSs. There is the possibility that either (a) a potential interaction exists between NS status and SLA education (at least for those who are NSs who have not had SLA education) or (b) the same finding for the two ID factors are actually one in the same, that those particular instructors simply corrected with more explicit feedback. The second possibility is more likely, especially in light of the fact that the NNS instructors corrected more implicitly and, when compared by SLA education, those with SLA education corrected explicitly. However, the current dataset cannot answer both questions and they must therefore be left for future investigations.

**Instructor Teaching Experience**

There was one significant finding concerning the third and final instructor ID in terms of type of feedback provided: less experienced instructors ignored errors significantly more than those with more teaching experience; there were no significant differences in their use of explicit or implicit feedback. The less experienced instructors corrected 59.9% of errors, compared to 76% of overall errors corrected by more experienced instructors. As the issue of amount of feedback has been discussed in the previous section, it will not be repeated here.
Research Question 2: Summary

In conclusion, the second research question found that instructors did differ in the amount of feedback they provided according only to NS/NNS status; no significant differences were observed for instructors when compared by SLA education or years of teaching experience. All three ID factors were related to the type of feedback provided.

Research Question 3: Consistency of Feedback Provision

RQ3: Do instructors of FL Spanish differ in the consistency (i.e., providing $x$ type of oral corrective feedback for $y$ errors) of their oral corrective feedback provision? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

For RQ3, the null hypotheses were rejected: instructors of FL Spanish do differ in the consistency of their feedback provision (see Table 25). For lexical, morphosyntactic, incorrect answers and use of the L1, there were significant results for all three instructor ID variables. For phonological errors, there were significant results for instructor NS/NNS status and teaching experience. Due to the lack of research investigating instructor feedback consistency, the discussion is restricted to the current study.

Lexical Errors

With respect to lexical errors, there were significant results for all three variables. NSs were consistent at the level of significance in their provision of explicit feedback, meaning that they were significantly more likely to provide explicit feedback for lexical errors than provide implicit feedback or provide no feedback at all. The same was true for instructors without SLA education, and both more and less experienced instructors. NNS
instructors and those with SLA education were not consistent in their treatment of lexical errors, meaning that they were not significantly more likely to provide a certain type of feedback for lexical errors.

Like the previous RQ, both NS instructors and those without SLA education had the same results. However, this time, instructors, when compared according to experience, also corrected lexical errors with explicit feedback. Thus, it could not only be a matter of an instructor being both a NS and someone without SLA education.

**Morphosyntactic Errors**

All instructors were consistent in their treatment of morphosyntactic errors in one of two ways: either they were significantly more likely to provide explicit feedback for morphosyntactic errors, or they were significantly more likely to ignore them. With respect to instructor NS/NNS status, NS instructors consistently provided explicit feedback, while NNSs were significantly more likely not to address morphological errors. With respect to instructor SLA education, + SLA instructors were significantly more likely not to address morphological errors while those without SLA education corrected explicitly. This time, instructors with less teaching experience were significantly more likely to not address morphosyntactic errors; those with more experience were at the other extreme: significantly more likely to address the errors with explicit feedback.

With respect to this error type, the same results are seen for the variables of NNS and + SLA education again, and this time the results are also the same for NS and - SLA education. However, for this error, those with less experience had the same result as the NNS and + SLA instructor variables, and the +experience variable had the same result as the NS and - SLA factors. What is interesting here is that 17 of the 22 NNSs are also less
experienced. While this could explain why there are the same findings whether the participants are examined in consideration of NS/NNS status, SLA education or experience, this is only true for morphosyntactic errors. For the other types of errors, these patterns are not present. It appears that there are factors contributing to the type of feedback instructors provide, other than error type, and that perhaps these factors are the reasons instructors differ in their treatment of morphosyntactic errors.

*Phonological Errors*

For the treatment of phonological errors, there were fewer significant results. NNS instructors were again consistent in the way they did not address phonological errors, as were instructors with less teaching experience. In fact, both NNS instructors and those with less teaching experience corrected the least amount of phonological errors (28.6% and 44.8%, respectively). NS instructors, those with more teaching experience were not consistent in their treatment of phonological errors, nor were instructors when considered in terms of SLA education.

Phonological errors are not as straightforward to deal with as lexical, morphosyntactic or even incorrect errors, particularly when it comes to vowels. One could expect NNSs to not be as sensitive about phonological errors, coming from the same L1 as the majority of the learners. The less experienced instructors might have ignored phonological errors for the same reason (as many less experienced instructors were also NNSs), or perhaps they did not correct phonology for an entirely separate yet plausible reason: out of all the errors to correct, one could argue that phonology is the most complex and perhaps for this reason the less experienced instructors did not correct phonological errors.
**Incorrect Answers**

Regarding incorrect answers, NS instructors, those without SLA education and those with more teaching experience were once again consistent in their explicit feedback provision. Instructors who had SLA education used significantly more implicit feedback. NNS instructors and those with less teaching experience were not consistent.

For these incorrect answers, the most notable finding is the difference according to instructor SLA education. The fact that + SLA education was the only factor related to consistent implicit feedback provision for incorrect errors is interesting. Again, it seems that SLA education is related to unique feedback behaviors for L2 instructors, most likely for the same reasons posited for the other error types: something in their SLA education made + SLA instructors correct in an implicit way, either for this type of error or due to the presence of one or more additional factors during that moment of interaction.

**Unsolicited Use of L1 (English)**

NNS instructors were consistent in the way they did not address the use of English with feedback, as were those instructors without SLA education; this was the first and only time these two factors shared the same result. Instructors with less experience were consistent in their provision of explicit feedback.

**Research Question 3: Summary**

The results for the third RQ, which investigated instructor consistency in feedback provision, were ironically the most inconsistent thus far in the discussion. Originally, when analyzing the treatment of lexical and morphosyntactic errors, it appeared that the factors NNS, + SLA and –experience (and on the other hand NS, - SLA and + experience) had similar results. However, this pattern was not present for the treatment of
phonological errors, incorrect answers, the use of L1 and was therefore disregarded from further consideration in the current study. Notably, in this RQ discussion, NNS instructors did not provide implicit or explicit feedback consistently for any type of error; the only significant results associated with NNS status was the way they ignored phonological, morphological and L1 use errors, most likely due to the fact that NNS instructors had a feedback/error ratio of 50%. NS instructors, on the other hand, had a feedback/error ratio of 74.7% and had several consistent trends: NSs were statistically more likely to provide explicit feedback for lexical and morphological errors, as well as for incorrect answers.

According to instructor SLA education, those without SLA education were consistent for their treatment of lexis, morphology and incorrect errors, which were all addressed consistently with explicit feedback; use of the L1 was consistently ignored. Instructors with SLA education were only consistent for their treatment of incorrect answers, when these instructors provided implicit feedback. For morphological errors, + SLA instructors were consistent in the way they ignored these errors. It seems that explicit feedback was the type of feedback used by - SLA instructors regardless of error type, and that, + SLA instructors, perhaps due to their education, used a variety of feedback techniques.

Finally, more experienced instructors had the same feedback patterns as the NS factor: consistent explicit feedback provision for lexical, morphological and incorrect answers. The less experienced instructors, on the other hand, were consistent in their provision of explicit feedback for lexical errors and the use of the L1; these same instructors consistently ignored phonological and morphological errors. Although outside
the scope of the current study, it would be interesting to see if instructors changed with respect to the type(s) of feedback they provide as they become more experienced. During the SR comments, many instructors discussed the way they started on one extreme with feedback provision (e.g., correcting everything explicitly), switched to the other end of the spectrum (e.g., providing only implicit feedback) and, over time, found what they described to be a “happy medium.” Perhaps these more experienced instructors found what they believe works best for their students’ learning (explicit feedback) and they consistently do what they know works best. Additionally, in the SRs, not one instructor commented on why they did or did not address an error solely based on the error type; there were always additional internal and external factors cited as motivation for feedback provision, or lack thereof.

Research Question 4: Beliefs and Practice

RQ4: Do instructors’ beliefs about how oral corrective feedback should be provided in the classroom correspond with how they actually provide feedback to learners? If not, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

For the fourth research question, the null hypotheses were rejected. Instructors’ beliefs and practices do correspond, at least in two instances: for + SLA instructors regarding the use of implicit feedback and for less experienced instructors regarding the use of both implicit and explicit feedback.
**Instructor NS/NNS Status**

With respect to NS/NNS status, there was not one significant correlation between any belief and corresponding practice (see Table 14 for a list of each belief topic and corresponding practice; see bolded items in Table 26). This indicates that with respect to the amount of feedback, importance of feedback, use of implicit and explicit feedback, and treatment of lexical, morphosyntactic and phonological errors, neither NS nor NNS instructors behaved in the classroom as they believed they would. Both said they would reserve corrective feedback for certain situations (errors that impeded communication or were linguistic targets) but they corrected quite often, 74.7% and 50% of the time, respectively. And while both NS and NNS instructors believed they used implicit feedback often and explicit feedback the least, NSs used explicit feedback much more than implicit (47.4% compared to 27.3%) and NNSs used implicit and explicit feedback in comparable amounts (23.6% for explicit feedback and 26.4% for implicit). NS and NNS instructors also did not behave as they thought they would regarding the treatment of certain error types. While Both NS and NNSs said they believed morphosyntactic errors were the most important to correct, followed by lexical errors and then phonological errors, their behaviors were quite different. NSs corrected lexical errors most often, followed by phonology, and they corrected morphosyntactic errors the least frequently. NNSs also corrected lexical errors the most, however NNS instructors differed from NSs in that they corrected morphosyntactic errors the second most frequently, followed by phonological errors.

On the surface it appears that instructors, when compared by NS and NNS status, really do not do what they say they do for most aspects of feedback provision considered
in this study. However, it could be the fact that the lexical errors instructors chose to correct could have, in their opinion, impeded communication more so than the morphosyntactic or phonological errors. What is interesting is that while both NS and NNS instructors corrected more lexical errors, NNS instructors corrected morphosyntactic errors the second most frequently (phonological errors third), while NS instructors corrected phonological errors the second most frequently followed by morphosyntactic errors. Could the same logic apply, that instructors were correcting errors that impeded communication? The qualitative comments provided during the SR protocols certainly point to that answer; not once in the SR protocols did the instructors reference a certain error type.

_Instructor SLA Education_

Regarding instructor SLA education, there was one significant relationship between beliefs and practice: + SLA instructors had corresponding beliefs and practices regarding the use of implicit feedback (see bolded items in Table 27). They believed they used it fairly often, and they did (28.7% of the time). However, when it came to explicit feedback, both + SLA and - SLA instructors were not accurate with their explicit feedback provision: although both believed they used explicit feedback least frequently, + SLA instructors provided it even more than implicit feedback (30.1% of the time for + SLA instructors and 44.6% for - SLA instructors). There were also no significant relationships for the beliefs and practices regarding the treatment of certain error types. Although both + SLA and - SLA instructors believed that morphosyntactic errors were the most important to correct, followed by lexical errors and then phonological errors (the same order found when comparing instructors according to NS/NNS status), both
instructors with and without SLA education, in practice, corrected differently. When compared by SLA education, all instructors corrected lexical errors the most, followed by morphosyntactic errors and then phonological errors, although in different amounts. - SLA instructors corrected more in each category (77.8% of lexical errors, 66.8% of morphosyntactic errors and 59.1% of phonological errors) compared to + SLA instructors (63.2%, 54.2% and 53.2%). Finally, there were no significant relationships between beliefs and practices of + SLA and - SLA instructors with respect to the amount of feedback. Although both + SLA and - SLA instructors believed feedback provision should be reserved for select errors that impede communication or were the linguistic targets of the lesson, they corrected very often; 58.8% and 70.4%, respectively.

Why is it that the + SLA instructors, when compared to the - SLA instructors, had the only significant relationship between beliefs and feedback? What does it mean that it was implicit feedback? First of all, one could again assume that the instructors’ SLA education played a role, either directly or indirectly, in shaping instructors’ beliefs and practices regarding feedback provision. As the + SLA variable was the only ID factor to use implicit feedback at a level of significance in this study (in relation to beliefs as well as consistently for incorrect answers), the data point to the influence of SLA education. This influence appears to be limited to implicit feedback, however, at least in terms of the relationship between beliefs and practices.

**Instructor Teaching Experience**

Less experienced instructors had corresponding beliefs and practices regarding use of both explicit and implicit feedback (see bolded items in Table 28). More experienced instructors’ beliefs and practices about the treatment of lexical errors were significantly
related, although inversely so: while more experienced instructors believed morphological errors were the most important to correct, they corrected lexical errors significantly more.

For the variable of instructor experience, it seems that the less experienced instructors were more mindful of their feedback practices, perhaps because it was a newer experience for them. The more experienced instructors, on the other hand, appeared to be farther removed from the process, placing emphasis on lexical errors in an opposite way than they said they would.

Research Question 4: Summary

The overwhelming dissociation between instructors’ beliefs and practices points to the complexity of feedback provision in the FL classroom. Some studies, although not focusing directly on feedback, have found congruency between instructors’ beliefs and practices (e.g., Chavez, 2006). Basturkmen et al. (2004), however, was the most specific study investigating instructors’ beliefs and practices about error correction and found results parallel to the current study: instructors often do not do what they say they do.

Basturkmen et al. (2004) explained the inconsistencies in their study positing that, when instructors are asked about their error correction beliefs in the abstract, they rely on their technical knowledge about how something (in this case, error correction) should be done; during classroom interaction, however, instructors rely on their practical knowledge (p. 267). The researchers posit that these inconsistencies would decrease over time as instructors become more experienced and automatize their in-class behaviors. However, in the current study, the less experienced instructors were more consistent in their treatment of certain errors (RQ3) and more consistent between reported beliefs and
practices (RQ4). Thus, while it seems plausible that the inconsistencies in the current study might be a result of instructors’ accessing their abstract belief systems at one point (what Basturkmen et al. refer to as ‘technical knowledge’) and relying on their practical knowledge at another, instructor experience cannot be cited as the reason behind them, at least not with respect to the current dataset. What can be said is that results from the current study demonstrate that, while instructors may report very similar beliefs about oral corrective feedback (with SDs less than 1.8, regardless of the ID factor or belief topic being considered), their practices differ considerably (with SDs in the 20-40 range throughout, regardless of ID factor or classroom behavior).

The researcher was initially concerned that the 35 instructors who participated in the SR protocols would report different beliefs than those 25 who did not, since SR protocols participants would have an additional chance to reflect on their classroom behavior and feedback processes (and perhaps this would influence their response on the Belief Questionnaire). However, upon further qualitative analysis, the answers between the two groups were remarkably consistent; for example, while those who participated in the SR had a mean of 1.78 for a question, those that did not participate in the SRs had a mean of 1.77 for the same question. It is possible that this is further evidence of instructors using these two knowledge systems – one during the classroom or classroom-related behavior (such as the SR protocols) and one during more abstract tasks, such as completing a belief questionnaire. More studies are needed to see if these differences are consistent in additional contexts and if these findings are in fact indicative of instructors utilizing two different knowledge systems, as posited by Basturkmen et al. (2004).
Research Question 5: Reasoning Behind Feedback Provision

RQ5: Do instructors of FL Spanish differ in their reasoning behind their feedback provision choices? If so, are these differences related to instructor NS/NNS status, SLA education, or years of teaching experience?

For the fifth research question, the null hypotheses were rejected: instructors do differ in their reasoning behind their feedback provision choices. Many of these differences were related to instructor NS/NNS status, SLA education or years of teaching experience. The reader is reminded that, for each of the ID factors, stimulated recall comments were first analyzed according to the percentage of errors that instructors reported as noticed, followed by calculations of the number of noticed errors that instructors consciously reflected upon (henceforth referred to as “reflected” or “conscious” errors) (Table 29). Finally, the SR comments were analyzed qualitatively as belonging to one of four categories: external factors cited for errors corrected with oral feedback, external factors cited for errors not corrected with feedback, internal factors cited for those errors addressed with feedback, and internal factors cited for errors that were not addressed with oral corrective feedback.

Comparing NS and NNS Instructors

Although NNS instructors reported noticing and reflecting upon more errors than the NS instructors (67% and 71.2% as compared to 59% and 51%), regressions revealed no significant quantitative differences between noticed and reflected errors according to the NS/NNS ID factors. The qualitative analyses of the factors cited by instructors, however, did reveal differences. Overall, NS instructors cited 11 factors that NNSs did
not mention, and NNS instructors cited 21 factors that were not mentioned by NSs (see bolded items in Tables 30 and 31).

The most notable differences were the external factors NNS instructors cited regarding physical and auditory cues they received from students. While both NSs and NNSs cited a student pause as an external factor motivating their feedback provision, only NNSs cited factors such as the fact that a student looked at them, a student looked at another student in confusion, or the fact that a student used a questioning tone. Another interesting and expected difference was the way NNS instructors commented on their own knowledge, as well. Three NNSs said they did not correct because they were preoccupied with their own knowledge in the target language and did not want to correct something they were not absolutely sure was incorrect.

There are several questions that arise from these differences: why is it that NS instructors reported noticing and reflecting a bit less than NNS instructors, and of particular interest, why did the NNS instructors report more factors than the NSs? What does it mean for their feedback provision? The first two questions could potentially be answered by the differences of the NS and NNS groups: there were more SLA-educated NNSs than SLA-educated NSs. Perhaps this difference in education was the underlying reason; this will be further discussed in the next section. The third question, what does reporting almost twice as many factors mean for the NNS feedback provision? Well, NNS instructors corrected 50% of errors, compared to 74.7% of errors addressed with feedback by NSs. It seems logical to presume that, as a result of taking so many more factors into account during class time, NNSs did not correct as much. This could also explain the higher rate of feedback provision for NSs; as they reported thinking about less
factors, they were able to correct more as they were not preoccupied with deciding whether or not to correct. A more in-depth look at the SR comments shed more light on these differences.

While four NS instructors cited their own NS status in explanation of their error correction (or lack thereof), three times as many NNS instructors cited their NNS status as reasons why they did or did not address errors with oral corrective feedback. Interestingly, NS instructors commented on how error correction would be easier for them if they were NNSs, and the NNSs often said the opposite.

The NS comments focused on their NS status as something that made error correction more difficult. One instructor said that, being a NS it was more difficult to understand the learners’ vowels:

*I really focus on what they're trying to say. And sometimes that takes up all of my thoughts and energy. Like sometimes it is REALLY hard for me to figure out what they're trying to say. It's like speaking with someone who speaks not a different language but it's an interlanguage.*

Another said that if he had the same shared experience as the NNS students, it would make it easier for him to understand the mistakes and decide from there what would be important to correct or not.

Several NNS instructors cited challenges with error correction that went along with their NNS status, including a lack of confidence in their abilities: “If I’m not absolutely sure something is wrong, I won’t correct the error.”

Another instructor cited that because of the shared experience of being an L2 learner of Spanish, students do not automatically respect your knowledge and that, as a
result, NNS instructors must be particularly careful to maintain a line between student and teacher relationships.

_I think that [me being a] NNS they’re like ‘oh, she’s gone through this before.’ I think that that distance [between instructor and student] isn’t naturally there so I need to put it there._

Finally, several NNS instructors felt their NNS status assisted them in their feedback provision, and how grateful she was to not have to teach her native language.

_[Being a NNS], I remember what it was like for me being a student. Because a lot of time if you’re a NS of a language there are a lot of things that you may not be able to explain because you just KNOW- or at least maybe that’s just me and English. So whenever [my students] have the por and para questions, I’m able to get that concept across because I was in the same boat. I remember thinking ‘why can’t there just be para, why do we need por, too’?_

It seems that NS and NNS instructors do differ in their reasoning behind their feedback decisions in the FL classroom, and that part of those differences are the way they see themselves as NS or NNS instructors, which is predicted in the literature (e.g., Benke & Medgyes, 2005; Lasagabaster & Sierra, 2005; Samimy & Brutt-Griffler, 1999). This is not to say that either NS or NNS instructors are superior in terms of feedback provision, only that this ID factor influences their interactional feedback decisions.

_Comparing Instructors With and Without SLA Education_

+ SLA instructors noticed 69.5% of errors and reflected on 66% of them; - SLA instructors noticed 62% and reflected on 65.1%. Instructors with SLA education cited many more unique factors than those without SLA education (19 as compared to 11); see
bolded items in Tables 32 and 33. It makes sense that those instructors educated about
SLA would take more factors into account than those who were not. One would presume
that having taken SLA courses and learning about the L2 acquisition process and the
myriad variables that affect it would have an effect on how instructors think and act in the
L2 classroom, and this is corroborated by instructor comments during the SR protocols.

Many instructors cited how their SLA education complicated their feedback
provision. When asked about a certain morphosyntactic error that was not corrected, a +
SLA instructor stated:

*Should I stop her? Should I not stop her? Here's the problem. What do you do
when it comes to that? I think [my SLA course] was amazing. I also think it made
my job as a language teacher harder. For example, what [the student] just did,
when I used to teach public school I would have interrupted him. But now, the
problem is that it's a conjugated verb. Sometimes all you can do is just confuse
the person. I think the last thing I should do to him, an extremely shy guy, the
oldest and the only guy, is to hit him over the head. And just because I correct him
doesn't mean he'd understand. If it was a regular grammar day, it depends on the
activity. In a focused grammar activity, I would correct him...*

Because of her SLA education, this instructor cited a fear of overcorrecting, and an
intense online process during class regarding whether or not to correct an error. Another
instructor with a master’s degree in SLA said that he relied on his own instinct since after
graduate school “*I was really confused [about what to do in the classroom].*”

Another instructor found SLA education to be helpful, and cited how SLA
education changed his feedback provision habits and actually made him correct less:
I think reading the [SLA] research has really helped me become more aware of what I’m doing in the classroom and how I’m correcting... I think it makes me correct less. And I’m basing this on being a high school teacher before knowing the SLA literature. You know when I was in high school I would make students a verb wheel... now I think WOW what a waste of my time! Now knowing the literature and knowing how students process it’s completely different for me.

Quotes such as these help to explain the reasoning behind the finding that those with SLA education corrected less than those without SLA education; they also lend support to the hypothesis that SLA education does potentially change an instructor’s in-classroom behavior. At the very least there is a relationship between SLA education and in-class practices; more research is needed to claim a directional relationship. In light of the large amount of reported ‘noticed’ and ‘reflected’ errors along with the greater number of factors cited for feedback decisions (compared to those - SLA instructors), including the exemplary comments cited above, the relationship between instructor SLA education and in-class feedback provision is an area that clearly warrants further investigation.

Comparing More and Less Experienced Instructors

Instructors of differing experience reported the most similar percentages of noticed and reflected errors. More experienced instructors noticed 65% and reflected on 56%; less experienced noticed 61%, reflected on 63%. However, the qualitative comments paint an entirely different picture. As reported in Tables 34 and 35, instructors with less teaching experience cited more than four times the factors they took into consideration as compared to the more experienced instructors (almost all of who reported they had automatized error correction). More experienced instructors cited 6
factors that less experienced did not report; less experienced instructors cited 33 factors that more experienced instructors did not.

The more experienced instructors cited factors such as their perception of student understanding or student intention, as well as a certain student’s learning style or cultural stereotype. Instructors with more experience also corrected errors that had been previously addressed or errors that, in their opinion (based on their previous teaching experience), needed feedback to be corrected.

The literature would predict that more experienced instructors would report less online processing than less experienced instructors, just as the results demonstrated. One of the most marked differences between more and less experienced instructors is believed to be the automatization of classroom behavior over time (e.g., Berliner, 1995, 2001). While this is a common finding in instructor expertise studies, it is difficult to compare the results from the current study with instructor experience studies. As existing studies have focused on teacher cognition in general during ESL classes (e.g., Gatbonton, 1999, 2008), there were no feedback comments from the SRs in those studies for comparison.

In the current study, many more experienced instructors explicitly stated that they had automatized feedback provision after trying various ways through trial and error. One instructor with 12 years of experienced said:

*Now [feedback provision] comes automatically. A few years ago I was thinking about it…. I think there’s been an evolution. I have gone from being very explicit to very implicit to finding a balance between the two.*

Another instructor with 35 years of experience stated that the majority of the feedback he provides is automatic, stating:
I’ve seen it all before…I pretty much know what’s coming. Every now and then a student has a mistake I didn’t foresee but other than that I don’t reflect much.

The less experienced instructors, on the other hand, reported 33 factors that their more experienced counterparts did not mention. The most frequently mentioned factor was a concern for student feelings and a desire to avoid anxiety. Student personality, age, and previous language experience were also mentioned frequently. With respect to external factors addressed with feedback, the less experienced instructors relied on the student’s look, tone or pause, or the fact that the student had already spoken (eliminating the concern for creating anxiety, according to two instructors). Additional factors cited included: course objectives; the instructor’s teaching methodology; the lesson focus; the error placement in the speech or lesson; a concern to maintain a good class dynamic; the fact that a student was sharing something personal; a desire for the student to work it out on their own, etc. Four less experienced instructors also mentioned their own competency or performance as instructors, something never mentioned by those more experienced.

For example, the less experienced instructors cited their inexperience as reason why they thought more during classroom interaction. One stated:

I haven’t taught much grammar so sometimes it takes me some time to decide how to provide and correct the information.

Another instructor stated a concern for how the failure to correct certain errors could negatively impact the entire class.

I also think about the repercussions of me correcting or not correcting. Because one time there was a student who made a mistake and then everyone else repeated it and it was because I didn’t correct it the first time.
From the number of factors cited and the representative comments stated above, it seems that less experienced instructors do have many more in-class thoughts about their feedback provision. This could be an explanation for the difference in amount of errors corrected by more (96%) and less (59.9%) experienced instructors.

*Research Question 5: Summary*

The data from the fifth and final research question provide insight into the complex nature of instructors’ in-class feedback decisions, why they correct some errors and not others, and how they believe their ID factors influence their feedback provision. These SR protocols were key to explaining why instructors did not differ in their beliefs about feedback but *did* differ significantly in their in-class feedback provision, and differed often in relation to their ID factors.

**Conclusions and Implications**

In light of the results discussed above and in the previous chapter, several conclusions can be drawn from the current study, as well as implications for SLA theory, methodology, and teacher education.

**Conclusions**

The current dissertation sought to investigate three instructor ID factors — NS/NNS status, SLA education, and years of teaching experience — in relation to feedback provision in the Spanish FL classroom. By triangulating quantitative and qualitative methods, the study demonstrated that instructors do differ in the amount and type of feedback they provide in the FL classroom, and that these differences are
significantly related to their ID factors. Data from stimulated recall protocols were able to explain some of the observed differences and shed light on the many internal and external factors instructors take into account when making their in-class feedback decisions. Finally, the study demonstrated that instructor beliefs often do not relate to their behaviors, and in some cases instructors do the opposite of what they believe they do.

Implications

There are several implications for SLA theory, methodology, and teacher education that emerge from the results discussed in the previous chapters.

Theory

The results from the current dissertation demonstrate that instructor ID factors are important to consider in their own right, and that SLA research can no longer focus only on the IDs of the language learner, as the majority of the research still does. While the interaction approach does recognize the interlocutor as the provider of feedback, and a few studies have investigated feedback provided by different interlocutors, this research has been limited in scope; the bulk of the research is and has focused on the feedback provided to the learner, and what the learner does with it. It is the researcher’s hope that the current study inspires more empirical research investigating instructor IDs in relation to feedback provision in and outside of the FL classroom. If the current results are found in replication and additional studies, the researcher would propose that, in time, the IDs of the interlocutor (particularly the instructor interlocutor) would be a necessary addition to current theoretical discussions of the benefits of interaction.
As articulated in the previous chapters, the focus of interactional feedback research has traditionally been on the feedback provision to the learner, how the learner perceives and uses the feedback, and what factors internal and external to the learner affect the facilitative affects of feedback (depicted in Figure 3 below).

*Figure 3. Model of Traditional Interactive Feedback Research*

As demonstrated in the literature review, several studies have investigated interlocutor factors that affect the provision and learner use of feedback (e.g., Mackey et al., 2004; Polio et al., 2006), as seen in Figure 4.
The current dissertation differs from these studies in its demonstration of how instructors’ feedback decisions are made from and informed by learner and instructor ID factors in addition to external factors shared by both the learner and instructor (the type of language classroom, the point in the lesson, the amount of time already spent on a current topic, etc.). Thus, interactive feedback provision would be a process more along the lines of what is depicted in Figure 5.
In this figure, interactional feedback is focused on not just from point when it is given to the learner, but all the way back to when the instructor interlocutor makes the decision whether or not to provide feedback, and the various factors that go into this thought process (or lack thereof, as seen with the more experienced instructors). Rather than
consider the interlocutor only in terms of the type of feedback provided, results from the current study would lend support that the instructor interlocutor be considered for their IDs, as well as the learner, and that investigations into the instructor and how their IDs affect the input, interaction, and opportunities for modified output that learners have, ultimately affecting the SLA process. Thus, the current model proposes the instructor interlocutor should be in a position of equal consideration alongside the language learner.

Even in studies not specifically focused on the language instructor, researchers must be more responsible in reporting information about the instructor and/or interlocutor providing feedback. Classroom studies must report as much detail as possible regarding the instructor’s background, if not more, than is presented about the learners. Likewise, laboratory studies must discuss the interlocutor background and not just stop at reporting NS/NNS status, which, as shown by the current study, is only one piece of the puzzle.

Methodology

In terms of methodology, the current dissertation demonstrates the importance of triangulating both quantitative and qualitative methods in studies investigating ID factors. Had only belief or behavior data been collected, the results would have been clear in very different ways: the belief questionnaires would have led us to believe instructors do not vary in their feedback use according to the ID factors of NS/NNS status, SLA education or years of experience, while the classroom recordings would have led us to believe that instructors differ greatly when compared according to IDs. By using both types of methodology in tandem, along with the stimulated recall protocols, this study was able to shed light on how similarities and differences in instructor beliefs and behavior according to IDs can be found in the same study.
Teacher Education

Perhaps the most immediate implication for the current dissertation is for teacher education. Instructors’ uncertainty and preoccupation with providing too much feedback to their learners was a concern voiced in almost all of the belief questionnaires and stimulated recall protocols, particularly by those instructors who did not have SLA education. Those instructors who did have SLA education felt, at times, overwhelmed by the amount of information they had regarding the efficacy of feedback. Teacher education courses must make sure to spend adequate time on feedback provision, and importantly, to discuss the recent findings in the literature and how to practically put those findings into everyday teaching behaviors. Language program directors and those in charge of teaching methodology courses should ensure there is ample time to discuss error correction so that, upon entering the L2 classroom, those instructors who are not SLA researchers will be comfortable using feedback provision and those who are SLA researchers do not feel overwhelmed with the information they have. After all, as one instructor stated on their belief questionnaire, “If [oral corrective feedback] doesn’t take place in the classroom, where will it?” Due to the limited input and interaction opportunities FL students have in their target language, maximizing their learning opportunities in the classroom is vital and, according to the interaction approach, oral corrective feedback plays an important role in that L2 learning.
Limitations and Future Research

The following sections detail several limitations that should be considered alongside the results, and outlines several areas for continued research on instructor ID factors in relation to feedback provision in the L2 classroom.

Limitations

The majority of the limitations in the current dissertation can be addressed in future research. In the current study, the videocameras used were only able to capture interaction that took place with the entire class; therefore instructor feedback during pair work was unfortunately not captured. This can be addressed in the future by using wireless microphones, minimally on the teacher and ideally on both the instructor and students. If this methodological adjustment were to be made, however, more than one class should be videotaped, as the students and instructor would not be used to having a constant physical reminder that videotaping is taking place attached to their person.

Two limitations with data coding can be addressed in future studies. First, student error and teacher feedback can be coded for linguistic target. Although outside of the scope of the current study, by coding if an error or feedback was focused on a linguistic target of the class or course in general, additional result interpretation would be possible. Future studies could likewise extend interpretations by coding for a combination of implicit and explicit feedback use, in addition to implicit and explicit feedback categories. In the fourth research question of the current study, although instructors were asked if they used implicit, explicit, or a combination of the two types of feedback most
frequently, correlations run were between the belief question and implicit or explicit feedback, as there was no “combination” of implicit and explicit feedback coded.

Some limitations were inherent with the methods chosen and were thus unavoidable. First, inherent in all questionnaire data, instructors may not be forthcoming with their actual beliefs and may not accurately report their true opinions. In this study, the use of SR protocols assisted in uncovering these beliefs by asking 35 of the 60 instructors face-to-face what they believe about feedback. SR protocols, however, also have limitations. As with all retrospective measures, what participants report may not be what they were actually thinking at the time of the original interaction, and there is also the issue that exposing participants twice to the interaction could influence their responses. However, as the participants in the current study were instructors who were teaching and interacting with students during the recorded lessons, concurrent introspective methods were not an option and thus the limitation is unavoidable.

Finally, there is always a potential in using dichotomies, particularly when it comes to ID factors such as those in the current study. While most would not argue against a dichotomy for native or nonnative speakers, a dichotomy is more of a concern when it comes to teaching experience. In the current study, this potential limitation was addressed in two ways: (1) by operationalizing teaching experience in three ways, as a continuous, categorical, and dichotomous variable based on the literature of SLA, teacher education and educational psychology as well as from a qualitative inductive analysis of the data; and (2) by ensuring that these three operationalizations were statistically
comparable by means of a correlation.\textsuperscript{22} Thus, every measure possible was taken ensure that the dichotomies used were the most appropriate and robust for the current dataset.

\textit{Directions for Future Research}

As this was the first study to investigate the three ID factors of instructor NS/NNS status, SLA education and teaching experience in relation to feedback provision in a single study, there are numerous ways to extend this line of research, utilizing the data already collected as well as with new data.

\textit{Future Research with Current Data}

First, the researcher plans on thoroughly utilizing the 60 hours of classroom recordings for additional analyses. Student production will be a main focus of subsequent analyses, which can examine opportunities for and student production of modified output, as well as analyses of student speech, such as fluency, accuracy, and complexity. Learner error and instructor feedback will also be coded for linguistic target of the lesson, to see if linguistic target is related to instructor feedback provision (as many claimed it was) and student use of feedback. The researcher also plans on examining the data according to instructor ID factors in relation to task use in the classroom, as well as in relation to additional task features to include task phase, focus, and task type. Some of the data in

\textsuperscript{22} Per the request of one of the committee members experience was also investigated via extreme group comparison, meaning that instructors either had less than 3 years of experience or more than 10; any participant with 3-10 years of experience was eliminated from this follow-up analysis. While only 25 of the 60 instructors were subsequently used for these analyses, which also revealed much lower statistical power, overall findings were statistically similar to those revealed for the entire dataset. The original dichotomy was therefore maintained for several reasons: (1) in order to be consistent with the rest of the analyses, which used simple dichotomies (NS/NNS status, +/- SLA education); (2) based on the sample size, eliminating 35 of the 60 instructors would have reduced statistical power beyond the recommended amount; (3) the three operationalizations of experience were statistically compared via correlational analysis and found to be statistically correlated, meaning that they tested the same thing and that the dichotomy used in this study was appropriate for the dataset; and (4) a key motivating factor in examining instructor experience in this dissertation was the desire to investigate instructors of varying years of experience– not just those who were new to teaching or who had been teaching for a long time– that to date have been ignored in the field.
the current set will be compared with research collected from five Portuguese foreign language instructors to see if the patterns found in the current study are present in another language context. Finally, the researcher would like to analyze the current data set, particularly responses from the questionnaires and stimulated recall protocols, from a sociocultural perspective to see what additional results could be found. In this dissertation the instructor is argued to be a key component in tandem to considerations of the learning context, a central factor in both cognitive and social theories in SLA, and it would be interesting to analyze the data in another way.

*Future Research with New Data*

In the future, the researcher hopes to collect the additional data needed in order to conduct analyses of interactions between instructor ID factors with respect to their beliefs, feedback provision, and reasoning behind their in-class feedback choices. If possible, the researcher will also collect the data needed to see if instructor gender is an ID factor of potential interest to interaction-based feedback research.

Ideally, this study would be replicated in various foreign and second language classrooms to see if these instructor ID patterns persist regardless of context. It would be particularly interesting to examine instructor IDs and feedback provision in settings where a more teacher-centered classroom dynamic remains due to culture.

Future studies would do well to consider student use of feedback by measures of modified output (proposed in the previous section), uptake charts or other, more traditional measurements commonly found in laboratory contexts: pre-, post-, and delayed post-tests. And now that, at least in the current study, instructor IDs have been found to relate to in-class feedback provision, it would be interesting to see if these
differences in feedback provision are present throughout a semester or longer, and eventually, what affect (if any) they have on short and long-term student learning.

Future Research: Summary

As demonstrated by the current dissertation, instructor ID factors are important to consider in their own right in addition to learner ID factors in the L2 classroom. Future explorations into instructor ID factors will provide insight into the complex nature of interactional feedback from a viewpoint too often overlooked — from that of the FL instructor — and help researchers understand the SLA process as thoroughly as possible.
References


Johnson, K. E. (1996). The vision versus the reality: The tensions of the TESOL


D. Freeman & J. C. Richards (Eds.), *Teacher learning in language teaching* (pp. 249–59). Cambridge: Cambridge University Press.


Schulz, R. A. (1996). Focus on form in the foreign language classroom: Students’ and


Appendix A

Invitation to Participate in a Research Study

Dear Colleague:

My name is Laura Gurzynski-Weiss and I am a Ph.D. candidate in the Department of Spanish & Portuguese at Georgetown. The purpose of this email is to invite you to participate in a research study exploring interaction in the Spanish language classroom. The study is part of my dissertation and will take place in the summer and fall of 2009.

Participation in the study requires no interference with your normal class time and will not require your students to participate in any way; my interest lies in your classroom interaction and your beliefs as a language instructor. Your participation would include completing two questionnaires, one at the end of this semester and one at the end of the fall semester, and allowing me to videotape one of your lessons. Some instructors will also be asked to spend twenty minutes talking with me shortly after following the taped lesson at a time convenient for you. If you are willing to participate in the study, the data collection procedure will be as follows:

1) **Background Questionnaire**: The first thing you would need to do is to complete a content form and background questionnaire online. The purpose of this questionnaire is to determine participant eligibility for the study. The questionnaire can be completed online at your convenience and should take no longer than 15 minutes of your time. The consent form can be accessed at this link: [http://www.surveymonkey.com/s.aspx?sm=S6_2ff_2bDwWO_2bwr3LSxE6EYDw_3d_3d](http://www.surveymonkey.com/s.aspx?sm=S6_2ff_2bDwWO_2bwr3LSxE6EYDw_3d_3d) and the Background Questionnaire at this link: [http://www.surveymonkey.com/s.aspx?sm=OU3iu3_2bocHNVL9UGFqA2Lg_3d_3d](http://www.surveymonkey.com/s.aspx?sm=OU3iu3_2bocHNVL9UGFqA2Lg_3d_3d) When you complete the consent form and click “Done,” you will automatically be taken to the Background Questionnaire, which you can complete immediately following the questionnaire, or at a time more convenient for you.

2) **Taped Classroom Lesson**: During the summer or the start of the fall semester, we will determine a convenient date for me to videotape one of your classes. I will then come to your class on the pre-determined day and set up a video camera on a tripod in the back of the room. Some instructors will also be asked to meet with me for 20 minutes shortly after the taped lesson at a time convenient for you.

3) **Instructor Beliefs Questionnaire**: The final step in the study is to complete a questionnaire on your teaching and language beliefs. This questionnaire will also be emailed to you and can be completed online at your convenience. It should take no longer than 15 minutes of your time.

In total, your participation would require no more than one hour of your time. The data will contribute to a better understanding of important issues regarding interaction in the foreign language classroom and will also allow me to collect data for my thesis (which cannot be completed without the data). The data will be kept confidential and at
no time will your identity be matched with the information. I will share the study results of the study with you in hopes that you might benefit directly from it, as well.

I can be reached via email (lkg6@georgetown.edu) or phone (202-687-6134) to answer any questions that you may have. If you are willing to participate in the study, please complete the consent form and background questionnaire at the above links at your earliest convenience. I will be contacting you during the next month to arrange a time to tape your class.

I sincerely thank you in advance for your time and consideration and I look forward to working with you on this exciting and important research project.

Best,
Laura Gurzynski-Weiss
Ph.D. Candidate, ABD
Department of Spanish & Portuguese
Georgetown University
## Participant Consent Form

### Georgetown University Consent to Participate in Research

**PROJECT TITLE:** Investigating Instructor-Student Interaction in the Spanish Foreign Language Classroom  
**RESEARCHER:** Laura Gurzynski-Weiss  
PhD Candidate. Department of Spanish and Portuguese

The Georgetown University Institutional Review Board (IRB) has approved this research project. For information on your rights as a research subject, call the IRB at (202) 687-1506. Reference # 2009-134.

### INTRODUCTION
You are invited to participate in this research study. Please take whatever time you need to discuss the study with the researcher. The decision to participate or not is yours. If you decide to participate, please type your name and the date in the space provided. Once you have completed this consent form, please click the link in the email provided to proceed to the Participant Background Questionnaire.

### BACKGROUND AND PURPOSE OF THE STUDY
The researcher is interested in examining interaction in the Spanish foreign language classroom. The researcher hopes to use what she learns to contribute to the growing body of knowledge in the field of foreign language instruction, and will analyze the data for her dissertation.

### GENERAL PLAN AND LENGTH OF STUDY
Participation in the study requires no interference with your normal class time and will not require your students to participate in any way. Your participation will include completing two questionnaires, one at the beginning of the study and one at the end, and allowing the researcher to videotape one of your lessons. Half of the participants will also be asked to spend half an hour talking with the researcher shortly after following the taped lesson at a time convenient for you. In total, this study will require between 30-60 minutes of your time outside of class (30 if you are not interviewed or 60 if you are).

### CONFIDENTIALITY
After the study is completed, the researcher will transcribe the interaction and the interviews. At no time will identifying information be matched with the data and comments in the study. All data will be kept in password-protected computer files or in a locked filing cabinet in the researcher's home.

### RIGHTS AND COMPENSATION
Your participation in this study is completely voluntary, and you will not be compensated for your participation. However, the researcher will be available to train participants in the research techniques used if you are interested, and the researcher will share all results and findings with you.
Participant Consent Form

PROBLEMS AND QUESTIONS
Please contact the researcher listed above if you have any additional questions.

RESEARCHERS’ STATEMENT
I have fully explained this study to the participant. I have discussed the procedures and treatments and have answered all of the questions that the participant has asked.

Signature of researcher: Laura Gurzynski-Weiss

Date: April 22, 2009

PARTICIPANT’S CONSENT
I have read the information provided in this Informed Consent Form. All my questions were answered to my satisfaction. I voluntarily agree to participate in this study.

Signature
Date
University
## Participant Background Questionnaire

### 1. Teaching Background

Please complete the questionnaire with as much detail as possible. All identifying information will remain confidential and will only be seen by the researcher. If you have any questions, please contact the researcher at laurenmelendo@gmail.com or 414-773-1619. Thank you for your participation! It is sincerely appreciated.

1. **How many years have you been teaching Spanish language classes?**

2. **What types of Spanish courses have you taught? Please check all that apply.**
   - Beginning
   - Intermediate
   - Advanced
   - Literature
   - Teaching Spanish
   - Spanish Linguistics
   - Spanish methods

3. **What Spanish language courses do you expect to teach during the fall 2009 semester? Please include language level, days of the week and the time each course will be taught.**

<table>
<thead>
<tr>
<th>Course, Level, Days, Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

4. **What role(s) do you think the teacher plays in the classroom setting?**

<table>
<thead>
<tr>
<th>Role(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Participant Background Questionnaire

5. Please rank your knowledge about the following language acquisition theories or approaches.

<table>
<thead>
<tr>
<th>Theory</th>
<th>1 (don’t know anything about it)</th>
<th>2 (have only heard of it)</th>
<th>3 (familiar with it)</th>
<th>4 (basic knowledge of it)</th>
<th>5 (know well)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicative Language Teaching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar-Translation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task-Based Language Teaching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction Approach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sococultural Theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audiolingualism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Have you taken courses or seminars on language teaching?

- Yes
- No

7. Please describe the course or seminar(s) in detail, including information about the topics covered, and information regarding where, when, and how long you attended.

8. Have you attended course(s) on language acquisition research or theory?

- Yes
- No

9. Please describe the course(s) in detail, including information about the topics covered, and information regarding where, when, and for how long you attended the course(s).

10. Please rank the influence of current research on your teaching.

- 1 (no influence)  - 2 (little influence)  - 3 (occasional influence)  - 4 (weekly influence)  - 5 (daily influence)
11. Please explain your answer to the previous question. How does current research influence (or not) your teaching? Please be as specific as possible.

12. Do you read any current empirical or pedagogical research?
   - Yes
   - No

13. Please list any empirical or pedagogical sources (e.g., journals, books, proceedings, etc.) that you read, and how often you read them.

<table>
<thead>
<tr>
<th>Source/ Frequency</th>
<th>Read</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source/ Frequency</td>
<td>Read</td>
</tr>
<tr>
<td>Source/ Frequency</td>
<td>Read</td>
</tr>
<tr>
<td>Source/ Frequency</td>
<td>Read</td>
</tr>
<tr>
<td>Source/ Frequency</td>
<td>Read</td>
</tr>
<tr>
<td>Source/ Frequency</td>
<td>Read</td>
</tr>
</tbody>
</table>

14. Have or do you conduct research on language acquisition or teaching?
   - Yes
   - No

15. If you answered yes to the previous question, please explain the project(s) you have worked on in as much detail as possible.

16. How would you define language acquisition?

17. Do you follow any language acquisition theory or teaching method?
   - Yes
   - No
18. If you answered yes to the previous question, please state which theory or approach(es) you follow in your class and how/why you made this choice.

19. If you answered "no" to question 17, please explain how/why you made this choice.

20. What internal and external factors do you believe affect the acquisition process?

21. Does your knowledge of language acquisition theory affect your teaching?
   - Yes
   - No

22. Please rank the influence of your knowledge of language theory and acquisition on your teaching.
   - 1 (no influence at all)
   - 2 (little influence)
   - 3 (occasional influence)
   - 4 (weekly influence)
   - 5 (daily influence)

23. Please explain your answer to the previous question. How does your knowledge of language theory and acquisition influence (or not influence) your teaching?
2. Language Background

24. Do you consider yourself to be a native speaker of Spanish?
   - Yes
   - No

25. How would you define "native speaker"?

26. What language(s) did you speak at home as a child?
   - Language, % of time spoken
   - Language, % of time spoken
   - Language, % of time spoken

27. What additional languages have you studied and for how long?
   - Language, years studied
   - Language, years studied
   - Language, years studied
   - Language, years studied
   - Language, years studied
   - Language, years studied
   - Language, years studied

28. What language(s) did you use at school?
   - Elementary/Primary School:
   - Middle School:
   - High School:
   - College:
   - Graduate School:
   - Other:

29. Do you believe being a native/non-native speaker influences your teaching?
   - Yes
   - No
30. Please explain your answer to the previous question, using specific examples if possible.
Participant Background Questionnaire

3. Contact Information

Thank you for filling out the background questionnaire. Please be sure to fill in your contact information so the researcher may contact you to schedule your participation in the study.

Thank you in advance for your time and I look forward to meeting you in the fall!

Best,
Laura Gurzynski-Weiss
Georgetown University
Department of Spanish & Portuguese

31. Please fill in your contact information so the researcher can contact you.

Name: 

University: 

Email Address: 


Appendix D

Stimulated Recall Questions

Instructions (to be read to the participant at the beginning of the session):
What we’re going to do now is watch a recording of the class you just had. I am going to stop the tape at various points of the lesson and ask you what you were thinking at that time during the class. If you do not remember what you thought at that time, please say so. If there is any time that you would like to stop the tape and comment on something, just let me know. Do you have any questions? Ready to begin?

For certain error episodes with or without feedback:

1. What do you remember thinking at this point during the class?

2. At that moment, did you notice the student error?

3. (If yes to #2): At that moment, what did you think of the error?

4. (If yes to #2): At that moment, why did you decide to (not) address the error with corrective feedback?
   a. At that moment, what made you choose that particular type of feedback for that error?
   b. If error not addressed with feedback: At that moment, what made you decide to not address that error with feedback?

5. Is there anything else you would like to say about this moment?

At end of the session—

1. Do you think the interaction moments I asked you about were representative of the lesson?

2. Do you think your use of feedback in this lesson is representative of your normal feedback provision? Please explain.

3. Is there anything else you would like to comment on about this lesson or your teaching style in general?

4. Tell me more about your teaching experience and background
   a. SLA courses? Pedagogical training?
   b. Years and type of language teaching experience?
Appendix E

Participant Beliefs Questionnaire

1. General Beliefs

Please complete the questionnaire with as much detail as possible. There are no “right” or “wrong” answers; please answer with your honest beliefs and opinions. All identifying information will remain confidential and will only be seen by the researcher. If you have any questions, please contact the researcher at . Thank you for your participation in the study!

1. What do you see as the main roles of being a foreign language instructor? Please rank in order of importance. Use each number only once.

<table>
<thead>
<tr>
<th>Least important</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>Most important (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Provide students with input in the foreign language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Provide students with oral corrective feedback (i.e., spoken feedback that addresses a learner’s error)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Motivate the students about learning a foreign language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provide students with cultural information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. Act as a reference for when students have questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Do you believe your individual characteristics (personality, native language, training, experience with language learning, etc.) influence your feedback provision?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

3. Please explain your answer to the previous question, citing specific examples if possible.
2. Beliefs about Oral Corrective Feedback

4. Do you believe you are consistent in the AMOUNT of feedback you provide to students?
   - Yes
   - No

5. How often do you generally provide oral corrective feedback? Please choose one.
   - Every error
   - Occasional errors
   - Only if the error impedes communication
   - Only if the error is a target item/structure for the lesson

6. Please explain your previous answer, including why you do what you do.

7. Do you believe you are consistent in the TIMING of feedback you provide to students?
   - Yes
   - No

8. When do you generally correct students? Please choose one.
   - Immediately after the error
   - After the student appears to be finished with the thought
   - After the student has finished speaking
   - Other (please specify)

9. In reference to the previous two questions, please explain WHY you correct students when you do (i.e., after the error vs. after they have finished speaking).

10. In what setting(s) do you generally correct students? Please rank in order of frequency.

<table>
<thead>
<tr>
<th>Setting</th>
<th>1 (least frequent)</th>
<th>2</th>
<th>3 (most frequent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>During whole class activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During pair activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-on-one with the student</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Participant Beliefs Questionnaire**

11. Does SETTING make a difference in the way you correct students? Please explain in detail.

12. Do you believe you are consistent in the TYPE of feedback (i.e., more or less explicit, overt correction, recasts, etc.) that you provide to students?
   - Yes
   - No

13. In general, HOW do you correct your students? Please rank in order of frequency. Use each number only once.

| EXPLICITLY: Informing the student of what they said was wrong, how it was wrong, or providing the student with the correct answer | 1 (least frequent) | 2 | 3 (most frequent) |
| IMPLICITLY: Implying that there is something wrong with what the student has said without telling them exactly what is wrong (i.e., by repeating the error in a questioning tone, repeating what they said but with the correct form, asking the student to repeat their answer, etc.) | | |
| COMBINATION of explicit and implicit feedback | | |

Other (please specify)

14. Please explain your answer to the previous question, including WHY you correct in that way.
15. Do you provide oral corrective feedback CONSCIOUSLY in the classroom?
   \(\checkmark\) Yes
   \(\checkmark\) No

16. If you answered "yes" to the previous question, please explain in detail what goes into your decisions. If you answered "no," please explain in detail what influences your provision of oral corrective feedback.

17. Do you take any EXTERNAL FACTORS (i.e., classroom size, language level, linguistic target of the class, etc.) into account when you correct your students?
   \(\checkmark\) Yes
   \(\checkmark\) No

18. Please explain your answer to the previous question, including WHICH external factors you take into account and HOW (or why you do not take any external factors into account).

19. Do you take any LEARNER FACTORS (i.e., learner’s anxiety level, abilities, personality, previous mistakes, etc.) into account when you correct your students?
   \(\checkmark\) Yes
   \(\checkmark\) No

20. Please explain your answer to the previous question, including WHICH learner factors you take into account and HOW (or why you do not take any learner factors into account).
**Participant Beliefs Questionnaire**

21. Please rank your agreement with the following statements:

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Instructors have the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>responsibility to correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mistakes uttered by</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>students.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. One of my major</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>roles as an instructor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>is to provide oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>corrective feedback.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. It is important to be</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>consistent with error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>correction.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. Oral feedback in the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>classroom is important.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Without oral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>feedback in the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>classroom, students</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>will continue to make</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the same mistakes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. Some errors are</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>more important to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>correct than others.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

22. Which errors are the most important to correct? Please choose each number only once.

<table>
<thead>
<tr>
<th></th>
<th>1 (least important to correct)</th>
<th>2</th>
<th>3</th>
<th>4 (most important to correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical errors (vocabulary)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syntactic errors (sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>structure/word order)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphological errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(gender/number/tense/subject-verb agreement)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(pronunciation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

23. Do you differ in the WAY you correct certain errors (i.e., lexical vs. phonological)?

<table>
<thead>
<tr>
<th></th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO</td>
</tr>
</tbody>
</table>
### Participant Beliefs Questionnaire

24. Please explain how you generally correct the following types of errors:

- **Lexical errors:**
- **Syntactic errors:**
- **Morphological errors:**
- **Phonological errors:**
- **Unsolicited use of the L1:**

25. What is your overall opinion of the use of oral corrective feedback in the classroom?
### Participant Beliefs Questionnaire

#### 3. Contact Information

26. Please add your contact information.

<table>
<thead>
<tr>
<th><strong>Name:</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>University:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Campus Address:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Address 2:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>City/Town:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>State:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>ZIP/Postal Code:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Email Address:</strong></td>
<td></td>
</tr>
</tbody>
</table>