THE PRO-DROP PARAMETER IN SECOND LANGUAGE ACQUISITION 
REVISITED: A DEVELOPMENTAL ACCOUNT 

Larry L. LaFond 

University of South Carolina
Acknowledgments

A great many people have contributed to this dissertation. I first seriously contemplated pursuing a doctoral degree in linguistics as a result of the extraordinary faculty of Old Dominion University in Norfolk, Virginia. Janet Bing, John Broderick, Carole Hines, and Charles Ruhl all lent so much support and encouragement to me, not only then, but also throughout the years that followed, that I found the courage to change careers and enter into the world of academia. My sincerest thanks go to each of them.

My years at the University of South Carolina brought me into contact with some outstanding scholars, mentors, and colleagues who have enriched my life in numerous ways. Of them, Rachel Hayes and Rakesh Bhatt deserve special mention. Rachel and Rakesh first sparked my interest in Optimality Theory and served as coauthors of two papers that spawned the larger concept pursued in this dissertation. Without their initial ideas and continued collaboration, this dissertation would have never been written. Thanks to Rakesh for remaining a member of my committee after joining the faculty of the University of Illinois at Urbana-Champaign.

Special thanks also go to Barbara Hancin-Bhatt, whose teaching first opened up to me the field of second language acquisition, and whose frank suggestions and astute guidance went into creation of this dissertation's initial proposal. Barbara's continued friendship and mentoring have been important for my academic and professional growth.

Laura Ahearn, Anne Bezuidenhout, and Ramona Lagos all graciously agreed to be members of this committee. All three brought to the dissertation a keen awareness of discoursal aspects of texts and conversations. Ramona provided me access to her classrooms to collect data and was herself important to me as a native speaker informant. Laura and Anne have both taught classes I have taken, and it was a result of these classes and my subsequent interactions with them that I grew to deeply appreciate their perceptive minds and kind spirits. Anne and Laura have always strongly supported me personally, and they have continually improved the quality of my work with their insightful comments.

Numerous others became involved in various aspects of the dissertation. I am grateful to John Grego of the University of South Carolina Statistics Laboratory for his help on test design and statistical analyses, and to Sean Barnette for his assistance in entering a massive amount of data into Excel. I deeply appreciated the efforts of my son, Gabriel LaFond, who labored intensively to help me construct the table of contents and other front matter, as well as the references for the dissertation. I am also indebted to a dear colleague from Old Dominion, Julie Moberly, who was kind enough to copy edit the entire dissertation once it was complete. Paul Boersma, Matt Ciscel, Don Cooper, Dorothy Disterheft, Stan Dubinsky, Kurt Goblirsch, Jane Grimshaw, Juana Líceras, Hyeson Park, Benjamin Slade, and Ana Teresa Pérez-Leroux all commented on certain aspects of my research during the last of couple years, and for their critiques, suggestions, and friendship, I am thankful.

Thanks to Lizette Laughlin, Roberto Irizarry, Beatriz Kellogg, Alexandra Rowe, and especially Carolyn Hansen (who assisted me in more ways than there is space to
mention here) for access to their classrooms to collect data. Special thanks go to all of my research subjects, both at the University of South Carolina and at the Pennsylvania State University. In this regard, I must also especially thank Jacqueline Toribio who was personally responsible for collecting all of the data from the Pennsylvania State University simply because I asked. Jacqueline's efforts doubled the number of informants for the grammaticality judgment task and greatly improved the dissertation.

My deepest gratitude goes to my dissertation adviser, Eric Holt, who somehow always found the time to patiently guide and collegially discuss every step of the dissertation process. Eric's knowledge, professionalism, and valuable suggestions not only assisted me with the dissertation — he also enthusiastically supported every aspect of my professional development in my final year of graduate school, providing important support for my teaching, presenting, publishing, and job search. Eric clearly demonstrated to me that excellence in academic research does not need to involve either an overblown ego or a sacrifice of one's humanity.

Finally, I wish to thank my closest friend — my wife, Deb. More than anyone, Deb was most directly affected by my decision to pursue this goal, and she endured the process with unusual grace. I owe what remains of my sanity to her, because she never allowed the pressures of the urgent to cause me to lose sight of the important. My dissertating year would have been far more painful had I not yielded to her wise demand that it include time to dance, recreate, and love. Her unwavering support and encouragement has blessed me more than I can express.
Abstract
The pro-drop parameter in second language acquisition revisited:
A developmental account

Larry LaFond

This dissertation applies a particular theory of language acquisition and representation, Optimality Theory (Prince and Smolensky 1993, Grimshaw 1997), and a particular learning algorithm within this theory, the Constraint Demotion Algorithm (Tesar and Smolensky 2000), to the problem of how second language acquisition of pro-drop takes place for learners whose first language does not instantiate the grammatical properties traditionally associated with pro-drop.

The overarching goal of this study is to provide an account of the developmental stages in the second language learning of three grammatical properties: null subjects, inversion, and *that-trace*. Although there is no lack of such accounts from earlier generative perspectives, the need remains for a comprehensive developmental account from an Optimality-theoretic perspective. This dissertation begins to address that need.

The study here is based on several empirical tests (a translation task, a pilot study, and a grammaticality judgment task) that were administered to 370 adult native English speakers studying Spanish at the University of South Carolina or the Pennsylvania State University. Each task was designed to investigate learner competencies regarding null subjects, inversion, and *that-trace*. A key conclusion from these studies is that the acquisition of Spanish by native speakers of English involves a reranking of universal syntactic and discoursal constraints in these languages. Specifically, this dissertation argues that acquisition of Spanish occurs through the demotion of certain syntactic constraints in the English native grammar so that these constraints are dominated by discoursal constraints in the Spanish second language grammar.

This cross-sectional study tracks learners through developmental stages, but it is also theory driven, because the theory of grammar used in this dissertation permits specific predictions about the interaction and relative importance of constraints in Spanish and English and, ultimately, of the acquisitional route learners take. The application of Optimality Theory to interactions between discourse and syntax in second language learning represents a new and potentially productive line of inquiry that may advance our understanding of both second language learning and grammatical theory.
# Table of Contents

Acknowledgments ........................................................................................................ ii
Abstract ......................................................................................................................... iv
List of Tables .................................................................................................................. vii
List of Tableaux .............................................................................................................. viii
List of Figures ................................................................................................................ ix
List of Constraints ......................................................................................................... x
List of Abbreviations ..................................................................................................... xi

Chapter 1. Introduction ................................................................................................. 1

Chapter 2. Theoretical Background (Pro-Drop) .......................................................... 9
  2.0 Introduction ............................................................................................................ 9
  2.1 Defining pro-drop ................................................................................................. 10
  2.2 Early proposals regarding pro-drop ..................................................................... 14
  2.3 Parameter-setting approaches ............................................................................ 19
  2.4 Later proposals regarding pro-drop ..................................................................... 22
  2.5 Minimalist accounts ........................................................................................... 28
  2.6 Optimality theory and pro-drop ......................................................................... 30
  2.7 A response to Newmeyer (2000) .................................................................... 44
    2.7.1 The restrictiveness problem ....................................................................... 49
    2.7.2 The functionality problem ....................................................................... 53
    2.7.3 Unmotivated competition ........................................................................ 55
    2.7.4 The syntax/discourse interface ................................................................ 66
  2.8 Limitations of the major proposals regarding pro-drop .................................... 71
Notes ............................................................................................................................ 76

Chapter 3. Theoretical Background II (Pro-drop in Acquisition) .............................. 80
  3.0 Introduction .......................................................................................................... 80
  3.1 Linguistic theory and SLA .................................................................................. 81
  3.2 The logical problem of language acquisition ..................................................... 87
  3.3 The Subset Principle and learnability .................................................................. 92
  3.4 Theory of access to UG and L1 transfer ............................................................. 99
  3.5 Pro-drop in SLA .................................................................................................. 104
  3.6 SLA accounts of pro-drop from and OT perspective ........................................ 116
Notes ............................................................................................................................ 122
### Chapter 4. Research Methodology

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0 Introduction</td>
<td>123</td>
</tr>
<tr>
<td>4.1 Research questions and hypotheses</td>
<td>124</td>
</tr>
<tr>
<td>4.2 The validity of grammaticality judgments</td>
<td>128</td>
</tr>
<tr>
<td>4.3 Research subjects and data collection procedures</td>
<td>132</td>
</tr>
<tr>
<td>4.4 Task items and predicted results</td>
<td>137</td>
</tr>
<tr>
<td>Notes</td>
<td>146</td>
</tr>
</tbody>
</table>

### Chapter 5. Results and Interpretation

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0 Introduction</td>
<td>148</td>
</tr>
<tr>
<td>5.1 Translation task</td>
<td>149</td>
</tr>
<tr>
<td>5.2 Pilot study</td>
<td>152</td>
</tr>
<tr>
<td>5.3 Grammaticality judgment task</td>
<td>155</td>
</tr>
<tr>
<td>5.3.1 Null subject results</td>
<td>158</td>
</tr>
<tr>
<td>5.3.2 Inversion results</td>
<td>163</td>
</tr>
<tr>
<td>5.3.3 <em>That-trace</em> results</td>
<td>167</td>
</tr>
<tr>
<td>5.4 Interpretation</td>
<td>172</td>
</tr>
<tr>
<td>5.5 Limitations</td>
<td>175</td>
</tr>
<tr>
<td>Notes</td>
<td>178</td>
</tr>
</tbody>
</table>

### Chapter 6. Analysis and Conclusions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 Introduction</td>
<td>180</td>
</tr>
<tr>
<td>6.1 Constraint demotion in stratified domination hierarchies</td>
<td>184</td>
</tr>
<tr>
<td>6.2 Null subjects</td>
<td>187</td>
</tr>
<tr>
<td>6.3 Inversion</td>
<td>198</td>
</tr>
<tr>
<td>6.4 <em>That-trace</em></td>
<td>202</td>
</tr>
<tr>
<td>6.4.1 An initial analysis of <em>that-trace</em></td>
<td>208</td>
</tr>
<tr>
<td>6.4.2 A revised analysis of <em>that-trace</em></td>
<td>211</td>
</tr>
<tr>
<td>6.5 Summary of the analysis</td>
<td>216</td>
</tr>
<tr>
<td>6.6 Conclusions and implications for future research</td>
<td>217</td>
</tr>
<tr>
<td>Notes</td>
<td>221</td>
</tr>
</tbody>
</table>

### Appendices

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A: Grammaticality Judgment Task (English)</td>
<td>224</td>
</tr>
<tr>
<td>Appendix B: Grammaticality Judgment Task (Spanish)</td>
<td>230</td>
</tr>
<tr>
<td>Appendix C: Informed Consent Form</td>
<td>236</td>
</tr>
</tbody>
</table>

### References

<table>
<thead>
<tr>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>238</td>
</tr>
</tbody>
</table>
List of Tables

Table 2.1 Typology of empty categories ........................................... 17
Table 2.2 Constraints in Grimshaw & Samek-Lodovici (1995) .......... 35
Table 2.3 Constraint rankings for English and Italian ..................... 35
Table 2.4 Constraints used in Speas (1996, 1997) .............................. 38
Table 2.5 ‘Inviolable principles’ and their hedges ............................. 39
Table 2.6 That-trace constraints in Grimshaw (1997) ......................... 40
Table 3.1 Parameter model of acquisition of pro-drop ....................... 102
Table 3.2 Constraints implicated in the initial surfacing of null topics ... 118
Table 3.3 Distribution of research subjects, by level .......................... 135
Table 4.1 Analysis of initial parameter estimates for translation task, by group ................................................................. 150
Table 4.2 Analysis of initial parameter estimates for gramm. judgment task, by group ................................................................. 157
Table 5.3 Percentage of choice for null subject, by item and proficiency group ................................................................. 159
Table 5.4 Percentage of choice for overt subject, by item and proficiency group ................................................................. 161
Table 5.5 Percentage of choices for inversion, without focus .............. 163
Table 5.6 Percentage of choices for inversion, with focus .................. 164
Table 5.7 Average percentage of selection for inversion, by condition ... 164
Table 5.8 Percentage of choices including complementizer (+que/+null vs. -que/+null) ......................................................... 168
Table 5.9 Percentage of choices including complementizer (+que/+null vs. +que/-null) ......................................................... 170
Table 5.10 L2 developmental order: null subjects, inversion, that-trace ... 175
List of Tableaux

Tableau 2.1 Sample tableau adapted from Samek-Lodovici (1996)...... 32
Tableau 2.2 English topic-referring subjects ........................................ 37
Tableau 2.3 Extraction of a subject ......................................................... 41
Tableau 2.4 Constraints on *that-trace* .................................................. 42
Tableau 2.5 Complementizer optionality ................................................. 60
Tableau 3.1 English grammatical system ............................................... 119
Tableau 3.2 Spanish grammatical system .............................................. 119
Tableau 3.3 Initial constraint demotion: SUBJECT below DROP TOPIC .... 120
Tableau 3.4 Interlanguage 1: PARSE » DROP TOPIC » SUBJECT .......... 120
Tableau 3.5 Mark-data pair .................................................................. 120
Tableau 3.6 Constraint demotion: PARSE below DROP TOPIC .............. 120
Tableau 6.1 Stratified hierarchy C_3 » \{C_4, C_5\} » C_6 ....................... 185
Tableau 6.2 PARSE and SUBJECT in Italian ........................................ 188
Tableau 6.3 PARSE and SUBJECT in Italian reconsidered ..................... 189
Tableau 6.4 PARSE and SUBJECT in English ........................................ 190
Tableau 6.5a English ranking: PARSE » SUBJECT » DROP TOPIC » AF-Rt
(Subject is a topic) ........................................................................ 192
Tableau 6.5b English ranking: PARSE » SUBJECT » DROP TOPIC » AF-Rt
(Subject is not a topic) .................................................................... 192
Tableau 6.6 Mark-data pair: Beginning grammar (null subjects) ......... 193
Tableau 6.7 Demotion of PARSE ............................................................... 193
Tableau 6.8 Ranking after the demotion of PARSE ............................... 194
Tableau 6.9 Demotion of SUBJECT ........................................................ 197
Tableau 6.10a Advanced grammar (subject is a topic) ......................... 197
Tableau 6.10b Advanced grammar (subject is a non-topic) ................. 198
Tableau 6.11 SUBJECT » ALIGNFOCUS-RIGHT (English) ..................... 199
Tableau 6.12 ALIGNFOCUS-RIGHT » SUBJECT (Italian) ....................... 200
Tableau 6.13 SUBJECT, ALIGNFOCUS-RIGHT (Spanish) ...................... 200
Tableau 6.14 Intermediate ranking (inversion) ...................................... 201
Tableau 6.15 Advanced grammar (inversion) ........................................ 201
Tableau 6.16 Declarative complementizer (English) ............................. 206
Tableau 6.17 Declarative complementizer (French) ............................. 206
Tableau 6.18 Simple relative clauses (English) ...................................... 206
Tableau 6.19 Simple relative clauses (French) ..................................... 207
Tableau 6.20 Complex relative clauses (English) ............................... 207
Tableau 6.21 Complex relative clauses (French) .................................. 207
Tableau 6.22 Mark-data pair: Advanced (that-trace) ......................... 209
Tableau 6.23 Demotion of TELEGRAPH .............................................. 210
Tableau 6.24   Ranking after the demotion of TELEGRAPH ........................ 210
Tableau 6.25a Complementizer ‘optionality’ in English [-sub] ..........  213
Tableau 6.25b Complementizer ‘optionality’ in English [+sub] ....  213
Tableau 6.26   Obligatory complementizer in Spanish [+sub] ....  214
Tableau 6.27a Complementizer absence in English [+sub] ..........  214
Tableau 6.27b Complementizer absence in English [-sub] ..........  214
Tableau 6.28   Mark-data pair: Advanced (that-trace) ..........  215
Tableau 6.29   Demotion of T-LEX-GOV .................................................... 215
Tableau 6.30   Ranking after the demotion of T-LEX-GOV ..........  215

List of Figures

Figure 1.1 Syntactic constraints ................................................. 7
Figure 1.2 Discoursal constraints .................................................. 7
Figure 2.1 Distribution and reference of PRO and pro .................... 26
Figure 2.2 Operation of an OT Grammar ....................................... 31
Figure 2.3 Differences in ranking of a hierarchy .......................... 33
Figure 2.4 Alternations to a constraint hierarchy .......................... 33
Figure 2.5 Non-pro-drop language: Weak agreement features ....... 66
Figure 2.6 Pro-drop language: Strong agreement features ............ 67
Figure 2.7 T-model of grammar ................................................... 77
Figure 3.1 Subset relationship ..................................................... 93
Figure 3.2 Intersecting relationship of subjects in Spanish and English 94
Figure 3.3 Discourse conditions in Spanish and English ............... 95
Figure 3.4 Possible relationships between access to UG and L1 transfer ... 101
Figure 6.1 Syntactic constraints ................................................. 182
Figure 6.2 Discoursal constraints ................................................. 182
Figure 6.3 Stratified domination hierarchy .................................. 185
Figure 6.4 Rankings yielding non-null/null subjects ...................... 187
Figure 6.5 Grammar in flux ......................................................... 195
Figure 6.6 Ranking SUBJECT and ALIGNFOCUS-RIGHT .............. 199
Figure 6.7 Ranking TELEGRAPH and LEFTEDGE(CP) ............... 205
Figure 6.8 Rankings yielding that-trace ....................................... 212
Figure 6.9 Constraint demotion and the total ranking ................... 216
### List of Constraints

<table>
<thead>
<tr>
<th>Constraints and their first appearance</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGR(X)</strong> (p. 38)</td>
<td>A tensed verb hosts spec-head agreement between an agreement feature x and a nominal constraint.</td>
</tr>
<tr>
<td><strong>ALIGNFOCUS-RIGHT</strong> (AF-RT) (p. 7)</td>
<td>Align the left edge of focus constituents with the right edge of a maximal projection.</td>
</tr>
<tr>
<td><strong>CONTROL</strong> (p. 38)</td>
<td>A null pronoun must be controlled in its control domain.</td>
</tr>
<tr>
<td><strong>DROPFOCUS</strong> (DROPF) (p. 49)</td>
<td>Leave arguments coreferent with the focus structurally unrealized.</td>
</tr>
<tr>
<td><strong>DROPSUBJECT</strong> (DROPS) (p. 186)</td>
<td>Requires covert realization of subjects.</td>
</tr>
<tr>
<td><strong>DROPTOPIC</strong> (DROPT) (p. 7)</td>
<td>Leave arguments coreferent with the topic structurally unrealized.</td>
</tr>
<tr>
<td><strong>DROPTOPICrel</strong> (DropTrel) (p. 37)</td>
<td>Arguments with topic antecedents must be realized only minimally.</td>
</tr>
<tr>
<td><strong>FAITH[SUB]</strong> (p. 7)</td>
<td>The output value of [SUB] (for ‘subordination’) must be the same as the input value.</td>
</tr>
<tr>
<td><strong>FREEPRONOUN</strong> (p. 37)</td>
<td>A pronoun must be free in its governing category.</td>
</tr>
<tr>
<td><strong>FULL INTERPRETATION</strong> (FULLINT) (p. 32)</td>
<td>Parse lexical conceptual structure.</td>
</tr>
<tr>
<td><strong>LEFTEDGE(CP)</strong> (LE(CP)) (p. 7)</td>
<td>The first (leftmost) pronounced word in CP must be the complementizer.</td>
</tr>
<tr>
<td><strong>MAX(PRO)</strong> (p. 38)</td>
<td>If pro occurs in the input, its output correspondent is pro.</td>
</tr>
<tr>
<td><strong>NO PHI-FEATURES</strong> (p. 38)</td>
<td>Avoid agreement features.</td>
</tr>
<tr>
<td><strong>PARSE</strong> (p. 7)</td>
<td>Parse input constituents.</td>
</tr>
<tr>
<td><strong>RECOVERABILITY</strong> (REC) (p. 204)</td>
<td>that requires that the semantic content of unpronounced elements be recoverable from the local context.</td>
</tr>
<tr>
<td><strong>SUBJECT</strong> (SUBJ) (p. 7)</td>
<td>The highest A-specifier in an extended projection must be filled.</td>
</tr>
<tr>
<td><strong>TELEGRAPH</strong> (TEL) (p. 202)</td>
<td>Do not realize function words</td>
</tr>
<tr>
<td><strong>T-GOV</strong> (p. 40)</td>
<td>Trace is governed.</td>
</tr>
<tr>
<td><strong>T-LEX-GOV</strong> (p. 40)</td>
<td>Trace is lexically governed.</td>
</tr>
</tbody>
</table>
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 3</td>
<td>first person, second person, third person</td>
</tr>
<tr>
<td>Ø</td>
<td>null element</td>
</tr>
<tr>
<td>*</td>
<td>ungrammatical form or constraint violation</td>
</tr>
<tr>
<td>!</td>
<td>fatal constraint violation</td>
</tr>
<tr>
<td>K</td>
<td>optimal candidate</td>
</tr>
<tr>
<td>»</td>
<td>dominance relationship</td>
</tr>
<tr>
<td>,</td>
<td>non-ranking</td>
</tr>
<tr>
<td>AGR</td>
<td>agreement element</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary element</td>
</tr>
<tr>
<td>C, CP</td>
<td>complementizer, complementizer phrase</td>
</tr>
<tr>
<td>CDA</td>
<td>Constraint Demotion Algorithm</td>
</tr>
<tr>
<td>CV</td>
<td>consonant-vowel</td>
</tr>
<tr>
<td>cl</td>
<td>clitic</td>
</tr>
<tr>
<td>D-structure</td>
<td>abstract level of syntactic structure, 'deep structure'</td>
</tr>
<tr>
<td>EVAL</td>
<td>Evaluator that selects the optimal candidate GEN produces</td>
</tr>
<tr>
<td>fut</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>Generator that creates a candidate set of potential outputs</td>
</tr>
<tr>
<td>GB</td>
<td>Government and Binding Theory</td>
</tr>
<tr>
<td>iff</td>
<td>‘if and only if’</td>
</tr>
<tr>
<td>imperf</td>
<td>imperfect</td>
</tr>
<tr>
<td>infin</td>
<td>infinitive</td>
</tr>
<tr>
<td>INFL or I, IP</td>
<td>verbal inflection of a clause, inflectional phrase</td>
</tr>
<tr>
<td>L1</td>
<td>first language</td>
</tr>
<tr>
<td>L2</td>
<td>second language</td>
</tr>
<tr>
<td>LF</td>
<td>Logical Form</td>
</tr>
<tr>
<td>MP</td>
<td>Minimalist Program</td>
</tr>
<tr>
<td>MUH</td>
<td>Morphological Uniformity Hypothesis</td>
</tr>
<tr>
<td>N, NP</td>
<td>noun, noun phrase</td>
</tr>
<tr>
<td>OT</td>
<td>Optimality Theory</td>
</tr>
<tr>
<td>part</td>
<td>participle</td>
</tr>
<tr>
<td>perf</td>
<td>perfect</td>
</tr>
<tr>
<td>PF</td>
<td>Phonetic Form</td>
</tr>
<tr>
<td>pst</td>
<td>past</td>
</tr>
<tr>
<td>pl</td>
<td>plural</td>
</tr>
<tr>
<td>PRO, pro</td>
<td>phonetically null elements, ‘empty categories’</td>
</tr>
<tr>
<td>prs</td>
<td>present</td>
</tr>
<tr>
<td>RIP</td>
<td>Robust Interpretive Parsing</td>
</tr>
<tr>
<td>S</td>
<td>sentence</td>
</tr>
<tr>
<td>S-structure</td>
<td>abstract level of syntactic structure, surface structure</td>
</tr>
<tr>
<td>sg</td>
<td>singular</td>
</tr>
<tr>
<td>SLA</td>
<td>second language acquisition</td>
</tr>
<tr>
<td>Spec</td>
<td>specifier</td>
</tr>
<tr>
<td>SUBJ</td>
<td>Subject</td>
</tr>
<tr>
<td>Symbol</td>
<td>Term</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>t</td>
<td>trace element</td>
</tr>
<tr>
<td>TNS</td>
<td>tense</td>
</tr>
<tr>
<td>UG</td>
<td>Universal Grammar</td>
</tr>
<tr>
<td>V, VP</td>
<td>verb, verb phrase</td>
</tr>
<tr>
<td>wh-</td>
<td>class of words or phrases beginning with <em>wh-</em> in English (e.g. <em>who</em>, <em>what</em>, <em>which</em>, <em>why</em>, <em>which ones</em>)</td>
</tr>
<tr>
<td>XP</td>
<td>full phrase projection</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

This dissertation examines the pro-drop parameter in second language acquisition (SLA), based on a study of the acquisition of Spanish by native speakers of English. It argues that despite extensive research on pro-drop, accounts of pro-drop based on standard generative grammatical theories still contain a number of inadequacies. This dissertation explores an alternate account by applying a particular theory of language acquisition and representation, Optimality Theory (Prince and Smolensky 1993, Grimshaw 1997), and a particular learning algorithm within this framework, the Constraint Demotion Algorithm (Tesar and Smolensky 2000), to address the pro-drop question from a developmental perspective.

The overarching goal of this study is to provide an account of the developmental stages in the second language learning of three grammatical properties: null subjects, inversion, and that-trace. Although there is no lack of such accounts from earlier generative perspectives, and although some individual properties associated with pro-drop have received accounts within an Optimality Theory framework, the need remains for a comprehensive developmental account from an Optimality-theoretic perspective of the grammatical properties traditionally associated with pro-drop. This dissertation begins to address that need.

The study here is based on several empirical tests (a translation task, a pilot study, and a grammaticality judgment task) that were administered to adult native English
speakers who were studying Spanish at the University of South Carolina or the Pennsylvania State University. Each task was designed to investigate learner competencies regarding null subjects, inversion, and *that-trace*. The translation task involved 124 participants from four proficiency levels (beginning, intermediate, advanced, and native), a pilot of a 36-item grammaticality judgment task was administered to 39 intermediate learners of Spanish and the grammaticality judgment task involved 207 participants from five proficiency levels (beginning, intermediate, advanced, near-native, and native).

A key conclusion from these studies is that the acquisition of Spanish (a pro-drop language) by native speakers of English (a non-pro-drop language) involves a reranking of universal syntactic and discoursal constraints in these languages. Specifically, this dissertation argues that acquisition of Spanish occurs here through the demotion of certain syntactic constraints in the English native (L1) grammar so that these constraints are dominated by discoursal constraints in the Spanish second language (L2) grammar.

Therefore, this cross-sectional study not only tracks learners through developmental stages, but it is also theory driven, because the theory of grammar used in this dissertation permits specific predictions about the interaction and relative importance of constraints in Spanish and English and, ultimately, of the acquisitional route learners take. The application of Optimality Theory to interactions between discourse and syntax in second language learning represents a new and potentially productive line of inquiry that may advance our understanding of both second language learning and grammatical theory.
An overview of the dissertation follows: Chapter 1 (this chapter) provides a general description of the goals pursued in this dissertation, as well as a brief description of the contents of each chapter. At the beginning of each chapter, a more informative introduction outlines the goals and organization of that chapter. Chapter 2 discusses the major proposals regarding pro-drop as a theoretical construct. The chapter first looks at the early proposals to account for the cross-linguistic distribution of null subjects through a rule-based application of universal principles of grammar. Then it covers later proposals regarding pro-drop, focusing on theories of licensing and identification, morphological uniformity and, more recently, Minimalist accounts based on the strength of features. Subsequently, Chapter 2 considers a significant shift in linguistic theorizing to an approach that has provided new tools for reconsidering the pro-drop question — Optimality Theory (OT). The chapter provides a brief introduction to the major tenets of OT and discusses how recent research in OT may challenge previous conceptions of a ‘pro-drop parameter’ and may demonstrate the need to consider interactions of constraints from various levels of grammar. Chapter 2 also addresses recent critiques of OT on the basis that certain OT analyses (such as the one in this dissertation) violate the principle of the autonomy of syntax. Finally, the chapter reviews the limitations of the proposals regarding pro-drop to date, emphasizing the need for further empirical evidence from acquisition studies and for a developmental analysis that reflects that evidence.

Chapter 3 introduces major issues involved in acquisition research related to the question of pro-drop. Key issues considered in this chapter include the logical problem of language acquisition (Chomsky 1981), the subset principle (Angluin 1978, Berwick 1985, Manzini and Wexler 1987), transfer, access to Universal Grammar (UG), and
issues of learnability. The last of these, learnability, is a central concern to any
developmental account, and this dissertation focuses on a particular learning algorithm
emerging from research in learnability: the Robust Interpretative Parsing/Constraint
Demotion Algorithm of Tesar and Smolensky (2000). Chapter 3 reviews a number of
studies that have considered these issues from a parameter-setting perspective and two
that have provided SLA accounts from an OT perspective: Park (2000) and LaFond,
Hayes, and Bhatt (2001).

Chapter 4 begins with a discussion of the specific research questions and
hypotheses pursued in this study, the motivation for using a grammaticality judgment
task as the centerpiece for this study, the limitations and validity of a study based on
grammaticality judgments, and a brief description of how the grammaticality judgment
task in this study was constructed (using a translation task) and refined (through a small
pilot study). This chapter then discusses the research pool used for the translation task,
the pilot, and the main study, and the precise procedures used for collecting the data.
Finally, this chapter discusses in greater detail the creation of the tasks and the predicted
results. Four hypotheses are presented in Chapter 4. The first predicts that an
implicational hierarchy will be found between the initial acceptability of null subjects,
inversion, and that-trace, supporting the conclusion of Liceras (1989), who claimed that
null subjects are acquired before inversion, and inversion is acquired before that-trace.
The second hypothesis proposes that this hierarchy requires further refinement because
acquisition here must distinguish ‘initial acceptability’ of these properties from their
‘correct use’. The third hypothesis regards the ultimate attainment of learners and
predicts that at least some L2 learners will converge on native-like usage of the
grammatical properties associated with pro-drop, but that this convergence will come only in the later stages of L2 acquisition and as a result of a sensitivity to the discoursal constraints of the L2. The final hypothesis is that the developmental path taken by L2 learners will find a natural interpretation as the interaction of discreet discoursal and syntactic constraints, and that this interaction will provide support for the operation of the Constraint Demotion Algorithm of Tesar and Smolensky (2000).

Chapter 5 presents the results obtained from the translation task, the pilot task, and the main grammaticality judgment task. It discusses the coding and data analysis procedures that were used for each task, and then it presents the statistical results of those tasks. Significant differences between the responses of the various proficiency levels are discussed, organized around the three grammatical properties of null subjects, inversion, and that-trace. An interpretation of these results in light of the first three hypotheses of the dissertation is given. The empirical results reported in Chapter 5 show that learners begin to drop subjects indiscriminately in the early stages of acquisition. Soon after this, they begin to accept complementizer + null subject sequences, but they do not choose these sequences as consistently as native speakers. For native speakers, choices regarding subject inversion in declarative sentences remains relatively free in cases where the inverted constituent has the discourse feature of focus; where focus is lacking, the non-inverted choice is preferred. Although early learners select inverted options less frequently than later learners, their choices for inversion also appears to be affected somewhat by discourse conditions. Learners eventually identify the correct discourse conditions related to the dropping of subjects, and in the later stages of acquisition, they appear to acquire native-like judgments related to that-trace.
These results suggest that the first hypothesis, regarding evidence for the implicational hierarchy of Liceras (1989) of null subjects > inversion > *that-trace*, finds little support in this study. In this study, the acceptance rate of *that-trace* items among beginners is higher than the acceptance rate of inversion. Nevertheless, while this study does not provide additional support for Liceras’ hierarchy, it also does not disconfirm it. This study provides evidence that, even for early learners, the choice of inversion is affected by discourse context, and a lower acceptance rate of inversion does not demonstrate that inversion is acquired after *that-trace*, if choices related to inversion are optional in the native grammar but the choice of a complementizer in the native grammar is obligatory.

This study confirms the second hypothesis that claimed ‘initial acceptability’ must be distinguished from ‘correct use’ for each of these grammatical properties. For example, learners here made judgments indicating an acceptance of null subjects or *that-trace* long before they had accurate control of the properties of the L2 grammar regarding these conditions. Also, the third hypothesis, that some L2 learners will converge on native-like usage of these properties as a result of sensitivity to discoursal constraints of the target language, is supported by these results. This study showed no statistically significant differences between near-native and native speakers, and there is ample indication both that the shifts in grammar leading to the near-native proficiency level involved increasing sensitivity to discoursal conditions and that the relationship between these conditions and the syntactic constraints in the grammar is important.

Chapter 5 concludes with a discussion of the limitations of the study revealed by the results. The discussion focuses on certain interpretational differences that appear
inherent in the use of grammaticality judgments. It also suggests further refinements that could be made to the translation task (regarding the total number of speakers translating each item) and the grammaticality judgment task (regarding the number of items used to test particular subconditions).

Chapter 6 concludes the dissertation with an analysis of the developmental path taken by L2 learners in this study. In so doing, it addresses and supports the final hypothesis of the dissertation, that the learning path taken here displays an interaction between discoursal and syntactic constraints under the operation of the Constraint Demotion Algorithm (CDA) of Tesar and Smolensky (2000). The analysis in Chapter 6 argues that the developmental stages displayed in this data can be explained as an interaction between the syntactic and discoursal constraints shown in Figures 1.1–1.2:

Figure 1.1 Syntactic constraints

a. SUBJECT (SUBJ): The highest A-specifier in an extended projection must be filled. Failed when a clause lacks a subject in the canonical position. (Grimshaw 1995)
b. PARSE: Parse input constituents. Failed when input elements are not overtly parsed in the output.
c. FAITH[SUB]: The output value of [SUB] (for ‘subordination’) must be the same as the input value. (Baković 1997)
d. T-Lex-Gov: Trace is lexically governed. (Grimshaw 1997)

Figure 1.2 Discoursal constraints

a. ALIGNFOCUS-RIGHT (AF-RT): Align the left edge of focus constituents with the right edge of a maximal projection. Failed by non-aligned foci. (Grimshaw and Samek-Lodovici 1995)
b. DROP(TOPIC) (DROP(T)): Leave arguments coreferent with the topic structurally unrealized. Failed by overt constituents which are coreferential with the topic. (Grimshaw and Samek-Lodovici 1995)

Chapter 6 provides a detailed description of how the mechanism of constraint demotion works in the acquisition of null subjects, inversion, and that-trace. The chapter
also suggests some non-parametric explanation of why these conditions tend to pattern together (e.g. why inversion effects frequently appear in those languages that permit null subjects), concluding that the various properties associated with pro-drop are acquired by second language learners in a particular developmental order that is predictable by the CDA. These results challenge traditional parameter-setting accounts of pro-drop by arguing that the grammatical properties associated with pro-drop are actually epiphenomena resulting from particular constraint rankings within a grammar, not from the switching of a parameter to a particular inviolable setting.

This dissertation also argues against accounts in which autonomous applications of syntax fail to admit interactions between syntax and discourse. The developmental path taken by the L2 learners in this study is best characterized in terms of the interaction between these two components of grammar. There is additional evidence that the interaction between syntax and discourse for L2 learners is not parameterized because lower-ranked constraints sometimes remain operative and higher-ranked constraints are sometimes violated. This result would not be expected in a model of grammar built on inviolable principles.

Further applications of OT to questions of second language learning hold a great deal of promise. The goal of this dissertation is not only to provide new insights into important interfaces between syntax and discourse in L2 acquisition, but also to highlight the need for this type of research and encourage others to conduct further investigations that will advance our understanding of second language acquisition.
Chapter 2
Theoretical Background (Pro-Drop)

2.0. Introduction

The goals of this chapter are (1) to provide a critical review of pro-drop as a theoretical construct that has commonly been used to explain and classify differences between languages, (2) to review how recent research in Optimality Theory (Prince and Smolensky 1993) challenges previous conceptions of a ‘pro-drop parameter’ and demonstrates the need to consider interactions of constraints from various levels of grammar, and (3) to address recent critiques that these Optimality Theoretic approaches violate the principle of the autonomy of syntax.¹

To achieve these goals, Section 2.1 begins with a discussion of the cluster of properties that have been traditionally subsumed under the pro-drop designation. Section 2.2 discusses some of the early proposals to account for the cross-linguistic distribution of null subjects through rule-based application of universal principles of grammar. Section 2.3 reviews the move within syntactic theory to situate the pro-drop construct within a parameter-setting framework (Chomsky 1981); this section also discusses the major pro-drop proposals that emerged from that framework. Section 2.4 considers later proposals regarding pro-drop, focusing on theories of identification and morphological uniformity. Section 2.5 then considers early Minimalist accounts of pro-drop during the beginning of the last decade. Section 2.6 reviews a significant shift in the linguistic
theorizing that provided new tools for reconsidering the pro-drop question—Optimality Theory. This section provides a brief introduction to the major tenets of Optimality Theory (OT) and reviews the OT studies that pertain to interfaces between syntax and discourse, particularly as these interfaces relate to pro-drop.

Since this dissertation represents the first developmental account to characterize the acquisition of several features of pro-drop through the interaction of discoursal and syntactic constraints within an OT framework, affirming the legitimacy of such an approach is crucial. For this reason, Section 2.7 reviews, and responds to, one recent critique (Newmeyer, 2000) of the use of OT to address issues of syntax, particularly when these approaches use constraints from numerous grammatical levels. Finally, having discussed the major proposals regarding pro-drop to date and having defended an OT-syntax approach to the pro-drop question, Section 2.8 reviews some of the limitations of the earlier proposals, highlighting especially the need for empirical evidence from acquisition studies, and for developmental analyses that reflect that evidence.

2.1. Defining pro-drop

Linguistic theorizing during the last two decades has often attempted to explain why some languages permit subjects of tensed clauses to be null, and other languages do not. For example, certain utterances that are permitted in languages such as Portuguese (2.1a), Italian (2.1b), and Spanish (2.1c) are ungrammatical in a language such as English (2.2).

(2.1.) a. *Pedro disse que viu o João.* (Portuguese)
   Pedro, say-3sg-pst that proi saw-3sg-pst João
   ‘Pedro said that he saw João.’
b. Ø piove.
   rain-3sg-prs
   ‘(It) is raining.’

   (Italian)

c. Ø Compramos unos recuerdos.
   buy-1pl-pst some souvenirs
   ‘We bought some souvenirs.’

   (Spanish)

(2.2.) a. *Pedro said that saw Juan.

b. *is raining.

c. *bought some souvenirs.

Since Chomsky (1981), it has often been thought that sentences such as those in 2.1 contain a phonetically empty, but structurally present, subject. Chomsky suggested that these tensed sentences contain an empty pronominal element, referred to as ‘pro’.

One difference between the languages in 2.1 and languages that do not that permit such utterances amounts to whether or not a language permits the use of this pro element in tensed clauses. Those that do permit its use have been labeled ‘pro-drop’ or ‘null subject’ languages, while those that do not have been labeled ‘non-pro-drop’ or ‘non-null subject’ languages.

Chomsky (1981) recognized that pro-drop languages typically share a number of characteristics that differentiate them from non-pro-drop languages. For example, as others had earlier noted (Jespersen 1924, Perlmutter 1971, Taraldsen 1978), many pro-drop languages have ‘rich’ inflectional systems. Chomsky argued that, since the recovery of the missing subject was critical, these ‘richer’ language systems had unique qualities related to agreement (AGR) and inflection (INFL) that permitted this recoverability.

In addition to rich agreement, pro-drop languages were thought to typically display a cluster of associated morphological and syntactic properties. The following
examples from the pro-drop languages of Italian (adapted from Chomsky 1981), and Spanish (adapted from Liceras and Díaz 1999), illustrate this cluster of properties. Examples 2.3a and 2.3b illustrate the prototypical omission of a subject pronoun in a tensed clause:

(2.3.) a. ∅ ho trovato il libro. (Italian)
   b. ∅ he encontrado el libro. (Spanish)

   have-1sg find-part-pst the book
   ‘I have found the book.’

Pro-drop languages also often permit inversion in declarative sentences, as in 2.4, where the subject follows the verb:

(2.4.) a. ∅ ha mangiato Giovanni. (Italian)
   b. ∅ ha comida Juan (Spanish)

   have-3sg eat-part-pst John
   ‘John has eaten.’

In addition, pro-drop languages permit apparent violations of that-trace. This refers to a rule that, at least in English, prohibits clauses containing both the complementizer that and a ‘trace’, a phonetically null element left behind when a syntactic element moves in a clause as, for example, in ‘that t will leave’ in 2.5, where extraction of the subject from the lower clause requires the absence the complementizer in English, but not in Italian or Spanish:

(2.5.) a. Chi i credi che t_i partirà? (Italian)
   b. ¿Quién, crees que t_i se irá? (Spanish)

   who_i think-2sg-prs that t_i leave-3sg-fut
   ‘Who do you think (*that) will leave?’
Still other properties sometimes associated with pro-drop languages include the allowance of certain utterances that cannot easily be given English equivalents, for example, those containing long wh-movement (2.6) and empty resumptive pronouns (2.7) (‘a’ examples are Italian; ‘b’ examples, Spanish):

(2.6.) a. *L’uomo, che mi domando chi ∅, abbia visto.

the man that I wonder who has-3sg-pst see-part-pst

*‘The man (x) such that I wonder who x saw.’

b. *El hombre, que me pregunto a quién ∅, había visto.

this is the man who I wonder who has-3sg-pst see-part-pst

*‘This is the man who I wonder who has-3sg-pst seen.’

(2.7.) a. *Ecco la ragazza, che mi domando chi credi che ∅, farla.

this is the girl who I wonder who think-3sg that did it

*‘This is the girl who I wonder who thinks that (she) did it.’

b. *Esta es la chica, que me pregunto quién cree que ∅, lo hizo.

this is the girl who I wonder who think-3sg that did it

*‘This is the girl who I wonder who thinks that (she) did it.’

Expletive subjects (2.8) are often also included as a property of pro-drop languages:

(2.8.) a. ∅ ha piovuto.

(ITalian)

∅ has-3sg-pst rain-part-pst

‘(It) has rained.’

b. ∅ ha llovido.

(Spanish)

∅ has-3sg-pst rain-part-pst

‘(It) has rained.’

As will be shown later, not all these of these characteristics cluster together in every language that permits some of them, and this has led to some disagreement among researchers over what should or should not be included in a definition of pro-drop; nevertheless, there is broad acceptance that these languages contain empty categories and have phonetically unrealized, but syntactically present, elements. Furthermore, most have come to view pro-drop as a deep unlearned generalization, a universal, but parameterized, principle of grammar.²
2.2 Early proposals regarding pro-drop

Empty categories, those assumed to be syntactically present but phonetically null, have played a major role in syntactic theorizing. Early debate on pro-drop focused on distinguishing the characteristics of various empty categories, especially the empty category called ‘little pro’ (Chomsky 1982). Taraldsen (1978) had argued that null subjects are all empty Noun Phrases (NPs) bound in S’ by subject-verb agreement, as in 2.9, with the missing subject permitted because of the ‘rich’ verbal inflectional system.3

\[
\text{(2.9.) } [\text{NPI } e] \text{ comen} \_i \_a \text{ las diez}. \\
\text{eat-3PL at ten o’clock} \\
\text{‘They eat at ten o’clock.’}
\]

Jaeggli (1980), Chomsky (1981), and Suñer (1982) agreed with Taraldsen (1978) on the importance of a rich inflectional system, too, but they argued that the null element was not an empty NP but rather an empty element (PRO) posited by Government and Binding Theory (GB, Chomsky 1981) for control constructions where a non-finite verb had a null subject.

These early accounts of pro-drop were largely rule-based. For example, Jaeggli (1980) and Chomsky (1981) start from the position that sentences found in a pro-drop language (2.10a-c) share with non-pro-drop languages a rule of affix movement such as (2.11) that attaches INFL to the first verbal element at the level of Phonetic Form (PF).4

\[
\text{(2.10.) } a. \text{ Gianni telefonerà.} \\
\text{‘Gianni will call.’} \\
\text{b. Telefonerà.} \\
\text{‘He will call.’} \\
\text{c. Telefonerà Gianni.} \\
\text{‘Gianni will call.’}
\]

\[
\text{(2.11.) INFL V} \rightarrow \text{V + INFL}
\]
The rule in 2.11 leaves no trace, and in non-pro-drop languages, this rule always applies in the phonological component of the grammar, producing the structure in 2.12, where NP is governed by INFL:

\[
(2.12.) \text{NP'} \text{INFL'} \text{VP}
\]

PRO does not appear in a structure such as 2.12 because the subject position is governed by INFL, and PRO cannot be governed. But in pro-drop languages, rule 2.11 can apply not only in the phonological component, but also, optionally, in the syntax. When this option is not used, a sentence such as 2.9a surfaces, with the structure of 2.12 and nominative Case assignment applying to the NP governed by INFL. When the option is used, and 2.11 applies in the syntax, then the structure becomes that of 2.13, where the subject NP is not governed by INFL and so must become filled by PRO.

\[
(2.13.) \text{NP'} [\text{VP V + INFL'} ...]
\]

According to Chomsky (1981), The 2.10c case adjoins the subject to the VP, leaving a spot to be filled by a dummy pronominal (e.g. *there* or *it* in English), but in pro-drop languages, the minimal dummy allowed is PRO, resulting in 2.14:

\[
(2.14.) [\text{PRO'} [\text{VP telefon + INFL'}] \text{Gianni'}]]
\]

The structure in 2.14 not only satisfies the requirement that PRO be ungoverned, but it also allows the postverbal subject to receive nominative Case due to its being governed by INFL, as posited by GB.
Soon after this initial proposal using ‘big’ PRO, Chomsky (1982) introduced the empty category of ‘pro’ as a pronominal anaphor, justified as a separate category for several reasons, including: (1) obligatory fronting of the verb in Spanish questions (cf. Torrego 1984) would place PRO in a position where it would be governed (and this is not possible); (2) the missing subject in tensed clauses cannot normally be arbitrary in reference (unlike PRO); and (3) positing a pro category would permit the elimination of the stipulation or parameter that the rule in 2.11 can apply in syntax in pro-drop languages.

GB thus distinguished a number of empty categories. For example, in 2.15 and 2.16, the ‘emptiness’ was thought to be due to movement (resulting in a trace):

(2.15.) Mary$_i$ was kissed $t_i$ under the apple tree.

(2.16.) Who$_i$ do you think $t_i$ finished the pizza?

Empty categories could also be distinguished in regards to their referential features: NP-traces were considered [+anaphoric] and [–pronominal], while variables (empty elements, such as $wh$-trace or traces left behind by quantifier raising, which must be A-bar bound by an operator) could be considered [–anaphoric] and [–pronominal].

In contrast, in 2.17 and 2.18, there was thought to be no movement; the empty element is the subject of a non-finite clause, a pronominal anaphor (PRO) that, in regard to GB’s principles of Binding Theory, must be both bound and free within its governing category [+pronominal and +anaphoric]. Since this is not possible, PRO must remain ungoverned.

(2.17.) Harold$_i$ wanted PRO$_i$ to see the doctor.
A final empty category was ‘little pro’ (introduced by Chomsky 1982), the category that is most significant to a discussion of pro-drop languages. This ‘little pro’ was the understood subject of finite verbs, as in 2.19 and 2.20. Unlike ‘big PRO’, this category has the full referential properties that a personal pronoun would have; it is a pronominal non-anaphor [+pronominal, -anaphoric], which is to say that under GB’s Binding Theory, it is free in its local domain, not bound.

\[(2.19.)\] pro b\(\text{aila}\) bien. (Spanish, from Jaeggli 1982)
\[
\begin{array}{c}
dance-3sg \text{ well} \\
\text{‘She dances well.’}
\end{array}
\]

\[(2.20.)\] pro verrat. (Italian, from Rizzi 1982)
\[
\begin{array}{c}
come-3\text{sg-fut} \\
\text{‘He will come.’}
\end{array}
\]

As we shall see, GB theory treats sentences such as 2.19 and 2.20 as having an empty subject (different from no subject at all), because in this theory it is an inviolable principle that sentences have subjects (Extended Projection Principle, Chomsky 1981), regardless of whether or not they are visible.

Given the different distributions of these empty categories, a typology emerges:

Table 2.1

<table>
<thead>
<tr>
<th>Category</th>
<th>Pronominal</th>
<th>Anaphoric</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRO</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pro</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>NP-trace</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Variable</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
In addition to the typology given in Table 2.1, these empty categories could be contrasted in regards to how each category is handled by the different subtheories of grammar within the GB framework (e.g. Case theory, Theta-Theory, and Binding Theory). For example, in a GB framework, Theta Theory provides the impetus for creating a structural position (even if not phonetically filled) to which a theta-role may be assigned. In the case of expletive elements that do not bear thematic roles, as is the case with the English *it* in 2.21, Theta Theory does not force an expletive element, so something else must.

(2.21.) *(It) seems Isabel is sick.*

This motivated the Extended Projection Principle (EPP, Chomsky 1981), which essentially states that every sentence must have a structural subject. The Spanish equivalent of 2.21, given in 2.22, does not overtly satisfy this principle, so the conclusion must be that EPP is satisfied covertly, by the existence of an empty category.

(2.22.) pro *parece que Isabel está enfermo.*

seems-3sg that Isabel is sick

‘It seems that Isabel is sick.’

Thus, much of the early work in pro-drop revolved around identifying how the application of universal principles (together with the application of any rules) would account for the distribution of empty elements.
2.3 Parameter-setting approaches

A primary goal within linguistic theory has been to explain how children are able to acquire the complex nuances that a grammar presents them. Since it appears that children’s linguistic competence goes beyond what the input they have received would make possible, innate knowledge of some principles of grammar has been assumed. In the conception of Chomsky’s Principles and Parameters Theory (1981), children innately possess a set of universal and inviolable principles, as well as a set of parameterized options that can have more than one setting, and pro-drop has been thought to be one such parameter. If this parameter is set positively, the child’s language will display the distinguishing characteristics of pro-drop languages. If set negatively, the child’s language will not display these characteristics. Either way, the child’s grammar must still adhere to the conditions or constraints set by the universal principles of grammar.

Early applications of the parameter-setting approach involved the creation of new principles and new parameters. For example, the analysis of Bouchard (1983) posited a ‘Principle of Lexicalization’ and introduced a new parameter regarding where nominals are assigned Case. Bouchard argued that nominals are lexical if and only if they contain person, number, gender, and Case at Phonetic Form (PF). According to Bouchard, null subject languages consist of a parameter that optionally delays the percolation of Case assigning features of INFL until the level of Logical Form (LF). If the nominal is first assigned Case at LF by INFL, the nominal will not appear at PF. Bouchard then theorizes that in non-pro-drop languages the Case-assigning features of INFL must percolate up before surface structure, requiring Case-marked, overtly lexicalized nominals.
In another early paper, Suñer (1982) maintained that PRO and Ø are base generated and that trace is derived. She claims that traces are governed in both Spanish and English, but that PRO differs in these two languages. According to Suñer, in English PRO is never governed or Case-marked, but in Spanish, government of PRO is correlated with [±] tense. In Spanish, PRO is ungoverned in nonfinite clauses 2.23, but governed in finite clauses 2.24.

(2.23.)  \textit{Paco quiere} \textit{PRO comer.}  \\
Paco wants-3sg to eat  \\
‘Paco wants to eat.’

(2.24.)  \textit{PRO} \textit{i comen} \textit{i a las diez.}  \\
eat-3pl at ten  \\
‘They eat at ten o’clock’

The critical argument of Suñer (1982) is that PRO, pro, and traces in Spanish may appear in finite and nonfinite clauses, but in order to maintain this, she concludes that the government of PRO in tensed clauses in Spanish is a result of the lack of a subject requirement. For this, Suñer (1982) requires an additional parameter — the [±] obligatory subject parameter.

Rizzi (1982) argued that the \textit{that-trace} effect is suspended in certain null subject languages because in these languages the subject may first postpose, adjoining to the VP, where it is properly governed by V, and then is extracted from this proposed position, represented in 2.25.

(2.25.) \textit{[CP chi]} \textit{credi}  \textit{[S' che t, verrà t]}  \\
\textit{who believes that he will come}  \\
‘Who believes he will come?’

20
While this idea requires a certain ordering of movement, it lends support to a claim that *that-trace* effects are not the result of a parameter, but rather a fortuitous outcome of a grammar where the possibility of inversion of subjects permits an otherwise illicit *that-trace* sequence.

Within the parameter-setting model, most generative accounts have addressed the pro-drop question by attempting to isolate (1) licensing conditions (i.e. what permits the appearance of the null subject in a sentence) and (2) identification processes (i.e. how the referent of the empty subject is semantically recovered. For example, Rizzi (1982) made the significant claim that null subject languages differ from non-null subject languages in that the former have verbal inflections that are specified with a [+pronoun] feature with clitic-like properties. The subject pronoun cliticizes onto the verb, permitting recovery of person and number features. This insight was pursued in subsequent works (e.g. Jaeggli 1982, Bouchard 1983, Rizzi 1986, Roberge 1990, et al.) where the structure was assumed to be such that subject clitics (cl) can be proper licensors of pro as in 2.26:

\[
(2.26.) \\
\begin{array}{c}
IP \\
\text{NP}_i \quad \text{INFL}' \\
\quad pro \quad \text{INFL} \quad \text{VP} \\
\quad \quad \text{AGR} \quad \text{TNS} \\
\quad \quad \quad \text{cl}_i
\end{array}
\]

In 2.26, the subject clitic governs the [NP, IP] position. If lexical material is not present in the NP position, pro fills the position, receiving its licensing by the clitic under AGR. This provides an additional way for pro to be licensed. As Roberge (1990) observes, this proposal allows the licensing of pro in some languages (e.g. Spanish and
Italian) to occur through rich agreement, while licensing for pro in languages with subject clitics could be licensed by the clitics themselves. This divides languages into two additional categories — those with clitics and those without — and permits the parameterized option to be whether or not a language permits clitics. A general Recoverability Principle within a theory of government would still need to apply to both, meaning that empty categories must be licensed, but licensing could then come about in a variety of ways.

2.4 Later proposals regarding pro-drop

For many later studies in pro-drop, identification became the core property of null subjects. Early explanations of pro-drop were framed in terms of agreement features governing the empty category. The simple thought was that where there is overt agreement, subjects can be dropped. This observation alone frequently makes correct predictions. For example, Kenstowicz (1989) observed that person inflection was necessary for licensing in Bani-Hassan Arabic. In this dialect, if a verbal form inflects for tense, but not person (as in the participial form), then subject pronouns must be overt, but if it inflects for both person and tense, then null subjects are allowed (as in the perfect). Since it is possible that it is not a single parameter within a language that can blindly determine whether subjects can or cannot be dropped throughout a language, Kenstowicz (1989) provides evidence for a claim that more than just licensing is needed.

Rizzi (1986) also made such a claim when he posited that pro must be both licensed and identified. Earlier work in pro-drop had made much of the presence of verbal agreement features as the important element in licensing the empty subject
category. It had been assumed that when agreement features were rich enough, AGR possessed special governing powers. But Rizzi (1986) noticed that there are contexts where pro appears when identification through rich agreement is not possible. For example, in the Italian sentences in 2.27, where the objects are empty, the internal argument position requires pro, raising the problem of how this element is properly identified.

(2.27.) a. *Questo conduce [la gente/Ø] alla seguente conclusione.*
   ‘This leads (people) to the following conclusion.’

   b. *Gianni è sempre pronto ad accontentare [la gente/Ø]*
   ‘Gianni is always ready to please (people).’

Rizzi’s solution was to differentiate licensing from identification. Whereas licensing simply refers to the sanctioning of a given constituent, identification refers to the way the interpretation of empty categories is determined when only implicit null subjects are used. One language may choose INFL as a licenser for pro in the subject position and V for pro in the object position; another language may prohibit licensing altogether. Thus, the class of licensing heads may vary cross-linguistically, from permitting every head to be a possible licenser to allowing no licensing heads.

In terms of identification, Rizzi argued that referential pro is allowed if INFL is specified for the agreement feature of person, but the identification procedure is optional. Thus, some languages (e.g. German) will have an INFL specified for person, but will still not allow referential null subjects. Such languages will not be pro-drop. In other languages (e.g. Spanish) the null subject will have a grammatical specification of features on its INFL. Thus, pro will be both properly licensed and identified.
In contrast to Chomsky (1981) and Jaeggli (1982), for whom identification required agreement with the phi-features\(^8\) in INFL (with licensing requiring Case and lack of government), in the proposal of Rizzi (1986), for identification to succeed, coindexation with rich agreement specification (or an extended notion of binding) was necessary with licensing involving government. Rizzi (1986:524) proposes the principle in 2.28:

(2.28.) pro is Case-marked by \(X_y\) where \(X\) is a governing head of the type \(y\).

As noted, for Rizzi (1986), identification takes place through \(f\)-features or rich agreement specification. This is shown in 2.29, where pro resides in the Spec of INFL governed by INFL, and identified by the rich agreement specification on the first verbal form, *hemos*, in INFL.

(2.29.)

```
INFL
  /\Spec I'
   |   pro
    I VP
     |     
    hemos; encontrado
    have    found
```

This analysis, still based on rich agreement features, partially works for languages such as Spanish and Italian, which have the option of a licensing head, but runs into difficulties with other languages such as Chinese or Korean, that lack agreement features but still permit null subjects.
The inadequacy of agreement alone, or even agreement and identification, to account for pro-drop cross-linguistically is taken up by Huang (1984, 1989) who demonstrated that not all pro-drop languages rely on overt agreement to recover the missing subjects. Huang (1984, 1989) demonstrates that Chinese, a language entirely lacking in agreement features, may drop subjects (and objects) from finite sentences, as in 2.30 and 2.31. According to Huang, the null arguments in 2.31 are optional.

(2.30.) \textit{Zhangsan kanjian Lisi le ma?}
\textit{Zhangsan see Lisi Asp Q}
‘Did Zhangsan see Lisi?’

(2.31.) \textit{(ta) kanjian (ta) le.}
he see him Perf
‘(He) saw (him).’

Since Chinese does not have overt agreement, the null subject in 2.31 is problematic for accounts that rely on local licensing and identification.\(^9\) Huang (1989:193) addresses this problem through a generalized control theory outlined in 2.32 with the notion of control domain in 2.33 defined in Manzini (1983) and Nishigauchi (1984):

\begin{quote}
\textit{(2.32.) Generalized Control Rule (GCR)}
An empty pronominal is controlled in its control domain (if it has one).
\end{quote}

\begin{quote}
\textit{(2.33.) \(\alpha\) is the control domain for \(\beta\) iff it is the minimal category that satisfies both (a) and (b):}
\begin{itemize}
  \item [a.] \(\alpha\) is the lowest S or NP that contains (i) \(\beta\), or (ii) the minimal maximal category containing \(\beta\).
  \item [b.] \(\alpha\) contains a SUBJECT accessible to \(\beta\).
\end{itemize}
\end{quote}
This rule specifies the environments in which pro and PRO (Huang treats both the same) must have ‘a local, unique, non-arbitrary antecedent’ (1989:194). Crucially, if the empty category does not have a control domain, then this type of control does not apply. In such a case, reference may involve long-distance antecedents, arbitrary reference, or even pragmatic considerations. Huang’s (1989) position is summarized in Figure 2.1:

Figure 2.1. Distribution and reference of PRO and pro

Is there a control domain for PRO/pro?

YES   NO

PRO/pro is allowed

Is PRO/pro controlled in that domain?

YES   NO

PRO/pro is allowed

PRO/pro is excluded (e.g. in object position, subject position in finite clause of English-type sentences, and in non-finite clause with a raising verb)

Figure 2.1 shows where PRO/pro is allowed according to Huang (1989). When it is allowed, its reference is either free (when there is no controller) or determined (when it has a control domain and is controlled in that domain). One important claim Huang (1989) makes is that where PRO is excluded, so is pro. He is thus able to conflate the
two categories PRO and pro into a single category, eliminating the need for Chomsky’s PRO Theorem (1981) that separately argued for an ungoverned PRO based on Binding Theory.

Jaeggli and Safir (1989) represent another analysis that attempted to account for the presence of pro-drop in both languages with rich inflectional systems and those lacking an inflectional system altogether. Jaeggli and Safir reformulated the pro-drop parameter with an argument that null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms. Jaeggli and Safir essentially claim that languages lacking morphological uniformity will not allow null subjects. Their definition of ‘morphological uniformity’ is given in 2.34:

(2.34.) Morphological Uniformity
An inflectional paradigm P in a language L is morphologically uniform iff P has either only underived inflectional forms or only derived inflectional forms.

The Morphological Uniformity Hypothesis (MUH) essentially claims that languages that lack morphological uniformity will not permit null subjects. For example, English and French lack morphological uniformity and also do not permit null subjects. In contrast, Spanish and Italian do exhibit (relatively) uniform inflectional paradigms in the richness of their morphology so they permit pro-drop; Chinese and Korean display uniformity by virtue of their complete lack of verbal inflections, so they also permit pro-drop. Thus, the MUH would predict that the only pro-drop languages are those that have this quality of a uniform paradigm.

One final claim often made in the early 1980s was that null subjects and free inversion were linked by the same parameter, or the same syntactic requirement (Chao
1980, Jaeggli 1980, Chomsky 1981, et al.). Chao (1980) and Safir (1985) challenged this claim based on certain Italian dialects that permitted inversion, but not null subjects, and Portuguese, which permits null subjects but not inversion. Although the details of the analyses of Chao (1980) and Safir (1985) are not critical here, the observation that inversion and the absence of null subjects are not necessarily conjoined points to the need for an analysis that treats these two phenomena as the result of conditions that could variably produce one or the other, or both, effects.

2.5 Minimalist accounts

During recent years, the rise of the Minimalist Program (MP, Chomsky 1995) as a dominant theory of UG has highlighted the core assumption that a radical economy in a grammar’s set of theoretical and descriptive apparatus is necessary. To some extent, the MP and OT both share a certain set of insights that have gained greater acceptance. First, both MP and OT depart from classical GB theory by introducing the notion that different derivations, mappings, and surface forms compete with each other. Particularly in OT, but to a degree also MP, the candidate that wins is the one that is ‘best’ in relation to the other candidates. Secondly, both MP and OT recognize that the syntactic component cannot be entirely autonomous — it can only be understood with reference to the lexical morphosyntactic and semantic systems of the grammar. These developments in grammatical theory permit new consideration of questions that have been resistant to explanation when viewed only on the level of syntax. Pro-drop is one of those questions.

Speas (1994) considered pro-drop from a Minimalist stance. Using Chomsky’s (1991) Principle of Economy (‘project XP only if XP has content’), Speas proposed that
what separates pro-drop from non-pro-drop languages is where affixes are generated. Some languages generate affixes in the syntax, others in the lexicon. If AGR features are attached to the verb, projection is not possible, resulting in a non-pro-drop language. If AGR features are not attached, that is, if there is a morpheme in AGR (or in the case of some languages, no AGR at all), then pro-drop becomes possible. In languages such as Chinese that do not have agreement morphology, there is no need for an AGRP projection, and ‘no need’ translates into no projection and no requirement for licensing conditions. For Speas (1994), the pro-drop parameter is a parameter based on whether inflection is syntactic or lexical.

In another Minimalist approach, Radford (1997) argues that languages differ in the strength of the agreement features carried by their finite verbs. He suggests that when finite verbs carry strong agreement features, nonauxiliary finite verbs can raise from V to INFL and they can have a null pro subject, but when verbs carry only weak agreement features, neither the raising nor the null subject is possible. Radford assumes that the strength of features is correlated to the richness of the agreement inflections, and that, in a language that has a rich system of agreement inflections, identification of the null subject is more easily recoverable. As an illustrative example, Radford uses the historical changes to the inflectional system of Old English (which he claims was pro-drop) that eventually yielded Modern English (which is not pro-drop).

Minimalist theory employs ‘feature checking’, a relation between two elements such that one or more designated features they share are eliminated, as in 2.35:

(2.35.) a. *When did you go?

b. *They went when.
In 2.35a the +wh feature of when is checked in Spec of CP against the +wh feature of C. If when or C can not check their +wh feature, the derivation crashes, as in 2.35b. Minimalism distinguishes between strong features, which must be checked in overt syntax, and weak features, which due to another principle (i.e. ‘Procrastinate’) must be checked in covert syntax.¹²

While research in an MP framework still recognizes some role for parameters, it attempts to reduce the sources of parametric variation, and it works for explanations that provide only a negligible role for learning in the domain of syntax, with a far greater role for learning in the domain of the lexicon. MP has not fully discarded its GB predecessor, however, and the well-formedness of constructions is still seen to be a function of language-independent universal principles, together with parameters revealed in the acquisition of the lexicon, such as whether a language has strong features or weak features.

2.6 Optimality Theory and pro-drop

This section introduces Optimality Theory by first discussing the basic tenets of this theory and then illustrating the theory through a discussion of OT research to date pertaining to the three aspects of pro-drop that are the focus of this dissertation (null subjects, inversion, and that-trace). Given the extraordinary number of papers written on pro-drop from a mainstream generativist perspective, this chapter’s review of that literature has been necessarily selective. There are far fewer OT treatments related to pro-drop, and the review here is a more complete representation of work done from that
perspective; however, three important papers are not included in this section: Baković (1997) will receive only brief mention in this section, but will also be discussed in a response to Newmeyer (2000) in Section 2.7; and two positions regarding the acquisition of null subjects (Park 2000 and LaFond, Hayes, and Bhatt 2000) will be reviewed in Chapter 3, which deals with second language acquisition.

Like MP and its predecessors, Optimality Theory (Prince & Smolensky 1993, Grimshaw 1997) is a theory of generative linguistics which assumes that humans are endowed with an innate capacity for language and a language-learning mechanism that exploits this capacity; however, unlike MP and its predecessors, OT proposes that what is universal in human languages is a shared but finite set of grammaticality constraints. More importantly, these constraints are all violable and potentially in conflict with one another, resulting in a hierarchy between the constraints that varies from language to language. In OT, therefore, language variation is not the result of different parametric settings, but the result of differences in the language-specific rankings of the same set of universal constraints. This means that the available candidates for the realization of a structure directly compete, and the candidate that wins is the one that is most harmonic, or optimal, with respect to the hierarchy of conflicting requirements in the grammar. Since grammars contain universal constraints that conflict with each other, the grammars' job is to manage this conflict. The mechanism involved in this management is graphically depicted in Figure 2.2:

Figure 2.2. Operation of an OT Grammar

```
Input  GEN  Candidate set  EVAL  Optimal candidate
```
As hearers receive input, a candidate generator (GEN) produces the set of all possible candidate structural descriptions of that input. A candidate evaluator (EVAL) takes this set of candidates and assesses them in light of the ranking of universal constraints particular to the language that is being learned. EVAL determines which candidates involve lesser or greater violations, given that particular language-specific ranking. The structural description that involves the least serious combination of violations is the optimal candidate. In such a system, there is always at least one candidate that is better than others, the candidate that has the fewest or least critical constraint violations.

Violations of constraints in OT are normally illustrated through the use of a ‘tableau’ (Tableau 2.1). Definitions of the constraints in Tableau 2.1 are unimportant for now; the tableau is used for illustrative purposes only. Constraints are ranked across the top, with the higher-ranked constraints to the left and the lower-ranked constraints to the right. Solid lines between two constraints indicate crucial rankings between those constraints; the use of a dashed line indicates that the constraints on either side are not ranked with respect to each other.

<table>
<thead>
<tr>
<th>Tableau 2.1  Sample tableau adapted from Samek-Lodovici (1996:43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&lt;\text{sing} (x), x=\text{John}, -, T=\text{pres.perf}&gt;)</td>
</tr>
<tr>
<td>a. K. (\text{he has sung})</td>
</tr>
<tr>
<td>b. (\text{has sung he})</td>
</tr>
<tr>
<td>c. __(\text{has sung})</td>
</tr>
</tbody>
</table>

Candidates are given in the left hand column, with the optimal candidate (that with the fewest and/or least serious violations) marked by use of the ‘K’ symbol. All violations
are marked with an asterisk (*), and an exclamation point (!) after an asterisk marks that violation as ‘fatal’, meaning that the constraint that is violated is one which eliminates the candidate from further consideration as an optimal candidate, due solely to the fact that another candidate exists which does not violate constraints as highly ranked as the one marked with (!). Areas that are shaded mark constraints that are no longer relevant, since a higher ranked violation has already eliminated the candidate. Finally, the top left column of the tableau may contain the input to the system, as it does here.

Thus, in OT acquiring a language does not mean learning the well-formedness constraints of these structural descriptions (UG provides that); rather, language acquisition involves learning the language-particular ranking of these constraints, from strongest to weakest. When these rankings vary between languages, as in Figure 2.3, acquiring the second language will involve some type of reranking.

Figure 2.3 Differences in the ranking of a hierarchy

Hearer’s Hierarchy   Constraint 1 » Constraint 2 » Constraint 3
Speaker’s Hierarchy  Constraint 2 » Constraint 1 » Constraint 3

Hutton (1996), in a discussion of sound change, provides a useful summary of the types of alterations to a constraint hierarchy that might occur 2.4:

2.4 Alterations to a constraint hierarchy  (adapted from Hutton 1996:4)

a. Promotion of constraints
b. Demotion of constraints
c. Creation of new connections between constraints (A,B → A»B)
d. Dissolution of connections between constraints (A»B → A,B)
e. Alteration of dominance relationship between two constraints (A»B → B»A)
As a theory of constraint interactions, OT differs significantly from the earlier parameter-setting framework. As stated above, in OT there are no inviolable principles, other than those that create the structure of the language-learning mechanism itself; there are also no binary parameters or conditions that depend upon the strength or weakness of features. Since violability of constraints is an essential feature of an OT grammar, constraints may be stated in their simplest version, without the use of ‘hedges’ to cover exceptions to what might be otherwise thought to be a universal principle.

One of the earliest OT syntax treatments, Grimshaw and Samek-Lodovici (1995), provided an analysis of null subjects and inversion (two pieces of the pro-drop ‘parameter’) that captured differences between Italian and English by looking at the interaction of syntactic and discoursal constraints governing topics and foci in these languages. Grimshaw and Samek-Lodovici demonstrated that referential pro-drop is restricted to topic constituents, and they also showed that the fact that subjects can appear post-verbally in Italian, but not English, also derives from the language-specific ordering of the same universal constraints for these two languages. Crucially, because of the approach that Grimshaw and Samek-Lodovici take, they are able to make robust claims regarding not only regarding which languages will or will not permit null subjects and free inversion, they are also able to make language-internal, construction-specific predictions based on the grammatical effects of the discourse status of arguments.

Grimshaw and Samek-Lodovici specify both some of the universal constraints governing subjects and discourse, and the rankings of these constraints in Italian and English (Tables 2.2 and 2.3):
Table 2.2  Constraints in Grimshaw & Samek-Lodovici (1995)

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUBJECT (SUBJ)</td>
<td>The highest A-specifier in an extended projection must be filled.</td>
</tr>
<tr>
<td>FULLINTERPRETATION (FULLINT)</td>
<td>Parse lexical conceptual structure.</td>
</tr>
<tr>
<td>DROPTOPIC (DROPT)</td>
<td>Leave arguments coreferent with the topic structurally unrealized.</td>
</tr>
<tr>
<td>ALIGNFOCUS (ALIGNF)</td>
<td>Align the left edge of focus constituents with the right edge of a maximal projection.</td>
</tr>
<tr>
<td>PARSE</td>
<td>Parse input constituents.</td>
</tr>
</tbody>
</table>

Table 2.3  Constraint rankings for English and Italian

<table>
<thead>
<tr>
<th>Language</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>PARSE » SUBJECT » FULLINTERPRETATION » DROPTOPIC » ALIGNFOCUS</td>
</tr>
<tr>
<td>Italian</td>
<td>FULLINTERPRETATION » DROPTOPIC » PARSE » ALIGNFOCUS » SUBJECT</td>
</tr>
</tbody>
</table>

Augmenting earlier research on null and inverted subjects, they do what previous accounts have not — they show the interface between syntactic and discoursal requirements. In so doing, they go beyond the claim that languages may typologically be divided into pro-drop and non-pro-drop languages to make the additional claim that it is possible to specify the environments where inversion and null-subjects will occur.

Grimshaw and Samek-Lodovici begin with the observation that in Italian foci are right-adjoined (2.36):

(2.36.) Q:  *Chi ha gridato?*  ‘Who screamed?’
A:  
   b.  *Ha gridato Gianni.
   c.  *Gianni ha gridato
   d.  *Ha gridato, Gianni.*
While it is true that English and Italian differ in that Italian allows postverbal subjects and English does not, it is not true that subjects in Italian 'freely' undergo inversion.

Grimshaw and Samek-Lodovici maintain that subjects can only appear (and must appear) at the right edge of the VP when they are contrastively focused. In Italian, they claim, the ALIGNFOCUS constraint dominates a SUBJECT constraint. Therefore, when a subject is contrastively focused, it must move from its canonical subject position to the right edge of the VP. In English, SUBJECT outranks ALIGNFOCUS. Therefore, the subject does not leave its canonical subject position, even when contrastively focused.

As Grimshaw and Samek-Lodovici note, when constraints related to discourse or information structure are allowed to interact with syntax, it becomes possible to see a particular instantiation of a sentence as an 'optimal form of a distinct input, related to the pragmatics of correction-contexts' (1995:597).

Grimshaw and Samek-Lodovici view OT as an improvement over parameter-setting frameworks that simultaneously over- and under-generate topic-dropping due to a formulation of the parameter that does not allow it to interact with different levels of language. By using interactions between violable constraints, Grimshaw and Samek-Lodovici's approach allows them to delimit the contexts where null subjects and focus-adjoining occur. Since the languages in question vary in terms of the dominance configuration between competing syntactic and discursal constraints, such an approach is needed to account for the data.

Samek-Lodovici (1996) extended the earlier work with Grimshaw, continuing with constraints involving a competition between discourse and syntax. Samek-Lodovici’s main concern was subjects and the constraints related to their surfacing in the
grammar. For Samek-Lodovici, null subjects appear due to unparsed overt subjects; they are not null in the input itself. Therefore, a null subject results only when a discoursal constraint requiring the dropping of topics dominates constraints requiring subjects to be structurally realized. This is seldom the case in English (which is a primary way in which English differs from pro-drop languages such as Italian or Spanish), but Samek-Lodovici concedes that DROP_TOPIC could be a gradient constraint, and suggests that another version of DROP_TOPIC might be DROP_TOPIC_rel, a constraint requiring that arguments with topic antecedents be realized only minimally. Under the assumption that constraints from differing levels of grammar (phonological stress, discourse, syntax) can and do compete with one another, Samek-Lodovici maintains that stressed and unstressed pronominals will incur a differing number of violations of the DROP_TOPIC_rel constraint when the input is a topic-referring subject (Tableau 2.2):

Tableau 2.2  English topic-referring subjects (from Samek-Lodovici 1996:56)

<table>
<thead>
<tr>
<th>&lt;sing (x), x=John_top, T=pres.perf&gt;</th>
<th>PARSE</th>
<th>DROP_TOPIC_rel</th>
<th>SUBJ</th>
<th>FULL_INT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. K  he has sung</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. HE has sung</td>
<td></td>
<td></td>
<td><em>!</em></td>
<td></td>
</tr>
<tr>
<td>c. __ has sung</td>
<td><em>!</em></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

In a pro-drop language such as Italian, the constraints would be ranked DROP_TOPIC_rel > PARSE > SUBJECT, and this order would still favor null subjects over structurally realized subjects.

In another account, Speas (1997) returned to the issue of null subjects using OT (recall that she had earlier (1994) provided an MP account of null subjects). Speas claims that all languages have unpronounced pronouns whose reference must be understood within a context. Whereas in previous accounts, the occurrence of these pronouns was
subject to licensing conditions, and the interpretation of them subject to identification conditions, in Speas (1997) the crosslinguistic facts are derived by the ranking of three constraints, CONTROL, FREEPRONOUN, and MAX(Pro). Definitions of these constraints are given in Table 2.4. The first two definitions are unhedged versions of principles in Principles and Parameters Theory; the third constraint is added by Speas.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL:</td>
<td>A null pronoun must be controlled in its control domain.</td>
</tr>
<tr>
<td>FREEPRONOUN:</td>
<td>A pronoun must be free in its governing category.</td>
</tr>
<tr>
<td>MAX(Pro):</td>
<td>If pro occurs in the input, its output correspondent is pro.</td>
</tr>
</tbody>
</table>

Speas argues that in languages such as Thai or Korean, which allow both subject and object pro, FREE PRONOUN and MAX(Pro) are higher ranked constraints than CONTROL. In languages such as Spanish or Mandarin, which have subject pro but not object pro, CONTROL and FREEPRONOUN are also higher ranked than MAX(Pro). But to explain differences between English and Spanish, Speas (1997) uses additional constraints (gleaned from Samek-Lodovici 1996), NO PHI-FEATURES (‘avoid agreement features’) and AGR(X) (‘a tensed verb should host spec-head agreement between an agreement feature x and a nominal constraint’).

Speas holds that the constraint rankings for English and Spanish are the same in regards to CONTROL, FREEPRONOUN, and MAX(Pro). The difference between them comes from the ranking of NO PHI-FEATURES and AGR(X), where English ranks NO PHI-FEATURES higher than AGR(X) and Spanish does the reverse.

It should also be noted that Speas (1997) provided an early defense of the use of OT for syntax by critiquing the claim that syntax is made up of inviolable principles. Her
strong argument was that every principle in Principles and Parameters Theory, while claiming to be universal, actually involves hedging. She illustrates this graphically by showing the principle, its essence, and its hedge (Table 2.5).

Table 2.5 ‘Inviolable principles’ and their hedges (adapted from Speas 1997:184)

<table>
<thead>
<tr>
<th>Principle</th>
<th>Essence</th>
<th>Hedge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended Projection</td>
<td>All clauses must have a subject...</td>
<td>...except for languages which lack overt expletives.</td>
</tr>
<tr>
<td>Principle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case Filter</td>
<td>An NP must have Case...</td>
<td>...unless it is null.</td>
</tr>
<tr>
<td>Principle A</td>
<td>An anaphor must be bound in its governing category...</td>
<td>...unless it is one of a special class of anaphors that need not be bound.</td>
</tr>
<tr>
<td>Principle B</td>
<td>A pronoun must be free in its governing category...</td>
<td>...unless it occurs in an idiom like lose her temper.</td>
</tr>
<tr>
<td>Principle C</td>
<td>A name must be free...</td>
<td>...unless its an epithet.</td>
</tr>
<tr>
<td>Empty Category Principle</td>
<td>A trace must be properly governed...</td>
<td>...where ‘proper government’ means government by a lexical head or by a close enough antecedent</td>
</tr>
<tr>
<td>Theta Criterion</td>
<td>All thematic roles must be assigned to an argument position, and all argument positions must receive a thematic role...</td>
<td>...except that the agent of a passive may be absorbed by the verb, and the thematic roles of nouns need not be syntactically realized.</td>
</tr>
</tbody>
</table>

In another account, Grimshaw (1997) provides the beginnings of an explanation of *that-trace* effects in English. To do so, Grimshaw uses two government constraints (Table 2.6):
Table 2.6 \emph{That-trace} constraints in Grimshaw (1997)

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-GOV</td>
<td>Trace is governed.</td>
</tr>
<tr>
<td>T-LEX-GOV</td>
<td>Trace is lexically governed.</td>
</tr>
</tbody>
</table>

T-LEX-GOV is violated whenever the trace is not governed by a lexical head. T-GOV is violated only if a trace is not governed by any head. Grimshaw notes her indebtedness to Déprez (1994) not only for T-LEX-GOV constraint, but also for the insight that English \emph{that-trace} sentences are ungrammatical due to an alternative, the \emph{that}-less sentence. Grimshaw’s remarks that the presence of ‘that’ does not prohibit the extraction of an adjunct or complement, but extraction of a subject is only possible when no ‘that’ is present.\footnote{14} Therefore, there is no problem with sentences such as those in 2.37 (Grimshaw’s example 46):

\begin{enumerate}
\item Who do you think (that) they will see t?
\item When do you think (that) they will see them t?
\end{enumerate}

Extraction of the object in 2.37a involves no violation of either constraint, whether the ‘that’ is present or not. Extraction of the adjunct in 2.37b involves no government at all, resulting in both constraints being violated, whether the ‘that’ is present or not. The result is that both options (with or without the ‘that’) incur the same number of constraint violations for the extraction of objects (here both incur no violations) and for the extraction of adjuncts (here both incur two violations). The result is true optionality.

The situation changes when there is extraction of a subject. Consider the sentences in 2.38a and 2.38b (Grimshaw’s example 47):
(2.38.) a. *Who, do you think that t₁ will see them?

   b.*Who, do you think that t₁ will see them?

This time the two candidates differ in the constraints they violate. *That* may govern the trace, but since the complementizer position is not a lexical category, C does not ‘lexically’ govern. Therefore the candidate that includes *that* in 2.40b violates the T-LEX-GOV constraint, while 2.40a incurs no violations. See Tableau 2.3:

<table>
<thead>
<tr>
<th></th>
<th>T-GOV</th>
<th>T-Lex-Gov</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>V [IP t₁ I [VP t₁ V ...]]</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>V[CP that [IP t₁ I [VP t₁ V ...]]]</td>
<td></td>
</tr>
</tbody>
</table>

Grimshaw’s analysis is also used by Baković (1997) who does not essentially disagree with Grimshaw (1997), but points out that in examples such as those given above, the relative ranking of T-GOV and T-LEX-GOV plays no role; Since T-GOV is satisfied by both lexical and nonlexical government, it is not a useful constraint for distinguishing forms.

In place of T-GOV, Baković substitutes the constraint FAITH[SUB], stating that the output value of [SUB] (for ‘subordination’) is the same as the input value. To explain what is meant by this, Baković assumes that the distinction between an embedded CP and an embedded IP is their specification for subordination (i.e. CPs are [+SUB] and IPs are [-SUB]). Baković further assumes that the specification of [SUB] is provided in the input, and that it is the constraint FAITH[SUB] that regulates differences between input and output in terms of the [SUB] feature. If this is the case, then the violations incurred by sentences in 2.39 would vary based on the input, illustrated in Tableau 2.4.
(2.39.) a. [Which coat]i do you know [CP that ti doesn’t fit?]

b. [Which coat]i do you know [IP ti doesn’t fit?]

Presumably, Baković’s constraints would yield different results if ranked differently in another language (e.g. Spanish), but his analysis does not show this.

Tableau 2.4. Constraints on that-trace (Baković 1997)

<table>
<thead>
<tr>
<th>Input: [+SUB]</th>
<th>T-LEX-Gov</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [Which coat]i do you know [CP that ti doesn’t fit?]</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>b. K [Which coat]i do you know [IP ti doesn’t fit?]</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: [-SUB]</th>
<th>T-LEX-Gov</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [Which coat]i do you know [CP that ti doesn’t fit?]</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b. K [Which coat]i do you know [IP ti doesn’t fit?]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, Costa (1997) analyzes a number of constraints that have been proposed in relation to focus and canonical word order. Costa observes that some languages (e.g. English) use intonation for their primary focusing strategy, while others (e.g. Portuguese) use word order. The examples in 2.40 demonstrate (using capital letters to indicate primary sentence stress) how intonation indicates focused constituents in emphatic sentences in English (adapted from Dalbor 1997:68):

(2.40.) a. ALEX will never do that!

b. Alex will NEVER do that!

c. Alex will never do THAT!

d. Oh, yes, Alex WILL do that!
This is at variance with a language like Spanish, where word order changes to show sentence stress in emphatic utterances, because in these utterances the element with sentence stress is placed on the right edge of the sentence 2.41:

(2.41.) a. ¡Nunca lo hará ALEJANDRO! (from Dalbor 1997:68)
   b. ¡Alejandro no lo hará NUNCA!
   c. ¡Alejandro nunca hará ESO!
   d. ¡Ah, sí, Alejandro lo HARÁ!

English syntactically requires an object NP, a direct complement of the verb, to appear adjacent to the verb, even when it is focused and would be more appropriately placed (from a discourse perspective) in the clause-final position. This differs also from Portuguese, where the object must move to a clause-final position when focused, ignoring syntactic requirements to be adjacent to the verb. This is exemplified in 2.42 (from Costa 1997):

(2.42.) a: Who does Paul know well?  a: Quem é que o Paulo conhece bem?
   b: Paul knows MARY well.  b: *O Paulo conhece a MARGARITA bem.
   b': *Paul knows well MARY.  b': O Paulo conhece bem a MARGARITA.

In the Portuguese version, 2.42b' is correct only when it is the object that is focused.

Costa (1997b) shows that a shift of focus to the adverb bem would reverse the acceptability of 2.42b and 2.42b'. This could occur, for example, after a question such as, Como é que o Paulo conhece a Margarita? 'How does Paulo know Margarita?'.

Costa (1997b) uses Portuguese, English, Dutch, and Icelandic to illustrate differences between languages where there is a correlation between word order and
discourse functions and languages where this correlation is missing. Costa observed that Portuguese subjects and Dutch and Icelandic objects are generated in a position where they are focused regardless of the need for Case licensing. English, in contrast, violates discourse-related constraints when an element needs Case. Costa’s observation challenges the Principles and Parameters framework (which would say that Case is an inviolable principle) while it provides strong evidence for discoursal and syntactic interactions of the same elements being resolved differently in differing grammatical systems.

2.7 A response to Newmeyer (2000)

The previous section introduced OT as a promising approach for dealing with interactions between discourse and syntax, pointing out that, although Optimality Theory (OT) is a generative approach (since it assumes learners possess an innate language acquisition device and that individual grammars are acquired as this device encounters linguistic input), this theory distinguishes itself from other generative approaches by viewing the universal principles of language as violable, and potentially conflicting, constraints.

OT was chosen as the theory of grammar that informs the analyses given in this dissertation for two chief reasons. First, OT provides a precise way to address ‘the problem of language acquisition’.

Current research (e.g. Tesar and Smolensky 2000, discussed in this section and in the next chapter) has directly focused on the question of ‘learnability’ and provided clear demonstrations that OT yields a learnable model of grammar. Second, the structure of OT permits constraints from various levels of
grammar (phonology, syntax, discourse, etc.) to vie for prominence in the determination of the grammaticality of an output.

Both of these avowed reasons for using OT are controversial. Mainstream generativists have challenged the claim that OT grammars are sufficiently restrictive to limit the class of possible languages. They contend that an OT grammar will ultimately fail to produce ‘all and only’ the grammatical choices in a given language, that it will produce ‘rogue grammars’, and that total hierarchies of ‘unmotivated competition’ result in grammars that are unlearnable. Furthermore, they have argued that using OT to address issues of syntax is an enterprise that is beleaguered by inherent and irresolvable problems, at the heart of which is the OT view that a grammar consists of a complete hierarchy of competing constraints. For critics, this view leads to a failure to respect the autonomy of different grammatical components, particularly ‘the autonomy of syntax’.

The position that different components of grammar must be analyzed independently from the rest of the grammatical system is at the center of a longstanding linguistic debate between formalists and functionalists. In this debate, formalist theories (e.g. Government and Binding, Principles and Parameters, Head-Driven Phrase Structure Grammar, Generalized Phrase Structure Grammar, Lexical-Functional Grammar, Minimalism, and in part Relational Grammar) have differentiated relationships between grammatical elements and sought to keep analyses fully separate from any semantic, pragmatic, or discoursal properties of those elements. In contrast, functionalist theories (e.g. the Competition Model, Role and Reference Grammar, Cognitive Grammar, Systemic Grammar, Emergent Grammar, Functional Grammar, and Construction Grammar) have largely rejected the compartmentalization of grammatical
form, insisting that the function of assigning and conveying meaning involves linking semantic and pragmatic elements to other formal parts of the grammatical system (e.g. syntax). Both camps agree that formalists are ignoring certain facts in the interest of articulating a theory; formalists embrace this idealization while functionalists seek ways to overcome it.

At least one common goal for both camps is to provide an account of the linguistic knowledge of native speakers of a language. Creating such an account is a complex task and has led the formalists to, as Culicover (1997:11) has noted, either ignore ‘certain facts that do not naturally fit’ or ‘decompose phenomena into components’ under the theory that it will be easier to deal with the components of language as isolated parts than as an undifferentiated whole. That the components of grammar are legitimately analyzed in isolation from one another has been the view of many since Chomsky (1975) sought to demonstrate that meaning plays little or no role in syntactic analysis. Chomsky’s infamous phrase, ‘Colorless green ideas sleep furiously’, illustrated the point. The fact that native speakers of English identify this sentence as grammatical, despite its semantic gibberish, is thought to demonstrate that syntax is fully separate from semantics.

Hale, Jeanne, and Platero (1977) also developed this idea of autonomy in their Modularity Thesis or Autonomous Systems view that held that generalizations regarding one component of grammatical structure are statable without reference to interpretation or use. In addition, Jackendoff (1983) articulated the same view, maintaining that the set of components for each level of the grammar (and here ‘grammar’ refers only to syntax,
semantics, morphology, or phonetics/phonology) have their own well-formedness conditions, primitives, or rules of combination.

It is easy to see that in a system where the autonomy of components is assumed, there is no place for the primitives or well-formedness conditions of one component of grammar to make any reference to aspects of representation of another component of grammar. Although correspondence rules may map representations at one level onto representations of another level, there is no sense in which a phonological or semantic (let alone a discoursal or pragmatic) rule might determine the syntactic acceptability or grammaticality of a structure.

As Newmeyer (1998) points out, the ‘autonomy of grammar’ argument itself decomposes into three independent hypotheses: first, that there is a system of primitive terms whose combination makes no reference to semantics, discourse, or system external factors (autonomy of syntax); second, that the language competence is not dependent on language use or performance (autonomy of knowledge of language); and third, that the cognitive system of grammar is distinct from other cognitive systems (autonomy of grammar as a cognitive system). Newmeyer (1998) assembles empirical evidence in favor of each of these hypotheses. To show that syntax is independent of discourse function, he looks at various syntactic principles (wh-movement, lexical government, SUBJ/AUX inversion, etc.) for which he claims there is no direct link between form and function. He shows, for example, that SUBJ/AUX inversion takes place in requests, offers, exclamations, and questions, such as those shown in 2.43a-d.
(2.43.) a. *Could you pass the peas, please?*

b. *May I help you?*

c. *Is that ever a beautiful sky!*

d. *Could South Carolina have a winning season?*

According to Newmeyer (1998), the fact that constructions such as these have no uniform semantic properties poses a challenge to semantic or pragmatic accounts. He concludes, ‘The grammatical properties of human language are best characterized in terms of autonomous formal systems’ (1998:365).

Newmeyer is correct in arguing that OT syntax violates the principle of syntactic autonomy; but the last 40 years of syntactic theorizing within mainstream generativist theory have demonstrated that analyzing grammatical components in isolation does not always result in more elegant or adequate explanations of linguistic data. Recent research even outside of OT (e.g. Liceras 1988, 1989, Liceras and Diaz 1995, Pérez-Leroux and Glass 1997, 1999, Pérez-Leroux, et al. 1999), has begun to recognize that learner grammars are not comprehensible without an examination of semantic and discoursal interfaces with syntax. Within OT, claims of interactions between levels of grammar are even more pronounced. The analysis given in Chapter 6 of this dissertation, operates from the assumption that there are simply too many facts that must be ignored if interface grammars are disallowed from the onset.

Newmeyer believes that the use of OT for purposes of syntactic analysis is ‘seductive’ because, while it demands precision (in that ranking constraints demands greater formalization than is found in some other generativist theories), it is also more flexible (in that constraints may be not only violable but also invented or ad hoc).
Nevertheless, Newmeyer believes this enticement should be resisted because OT syntax suffers from three inherent flaws: (1) The restrictiveness problem; (2) The functionality problem; and (3) The problem of unmotivated competition. Since the SLA program described in this dissertation is based upon an OT syntax approach, each of these arguments must be addressed.

2.7.1 The restrictiveness problem

Newmeyer does not believe that OT grammars are able to exclude logically possible, but naturally nonoccurring, language types, regarding as baseless the claims in many OT papers that free constraint ranking is capable of explaining typological distributions of grammatical elements. He believes that such claims can only be made because OT researchers presuppose that the only relevant constraints are the ones being addressed in that specific paper. As evidence for his critique, he argues that Grimshaw and Samek-Lodovici’s (1998) proposal that languages may have null topics as subjects, but not null contrastively focused subjects, falls apart if one adds a constraint he calls DROPFOCUS to the DROPTOPIC constraint used by Grimshaw and Samek-Lodovici. He defines this constraint as in 2.44 (Newmeyer 2000:11):

(2.44.) DROPFOCUS: Leave arguments coreferent with the focus structurally unrealized. [Failed by overt constituents which are coreferential with the focus].

Newmeyer believes such a constraint must be valid because in some languages focus can be pragmatically retrievable. As an example, he presents the following, where an
exchange such as 2.45 in English could be communicated via a dropped focus constituent in Portuguese 2.46 (Newmeyer 2000:12):

(2.45.) Q: *Who broke the plate?*
A: *You know who!*

(2.46.) Q: *Quem é que partiu o prato?* ‘Who is it that broke the plate?’
A: *Partiu o prato?* ‘Broke the plate?’

Newmeyer also suggests that given a particular question intonation, this type of ‘focus-dropping’ could also be possible in Serbo-Croatian. Newmeyer concludes, ‘Grimshaw and Samek-Lodovici are incorrect in claiming both that they (null focused subjects) don’t occur and, given the architecture of OT, that they can’t occur’ (2000:14). He further asserts the following:

One is never on safe ground in claiming that one’s particular OT analysis is ‘restrictive’ in any interesting sense of the term. The limited set of constraints exhibited in the tableau in a particular analysis might well seem to conspire to predict the impossibility of some typologically-dispreferred structure. But since every tableau (when fully expanded) contains every constraint, it seems inevitable that *some* constraint will always lurk in the background which, if ranked highly in some particular grammar, will lead to the generation of that dispreferred structure.

(Newmeyer 2000:14)

If this critique stood, it would clearly be damaging to the OT program; however, there are several sensible ways to respond to this criticism. First, one must surely question whether the interpretation that Newmeyer wishes to place on the exchange in 2.48 is syntactically encoded. OT does not claim to account for all idiomatic, non-productive interpretations that are highly context-dependent; neither does any other grammatical theory to date. If Newmeyer wishes to create a unique set of discourse conditions, with a particular type of question intonation, together with a great deal of shared contextual knowledge between the participants, and then argue for a particular
grammatically encoded interpretation, he is claiming far more than any OT analysis to date has done. Second, Newmeyer is surely correct that restrictiveness is a serious problem if constraints can be constructed willy-nilly in the manner of his DROPFOCUS example. Dropping focused constituents is not a general or motivated constraint for any language, including Portuguese and Serbo-Croatian. In contrast, the dropping of topic subjects is a well-attested phenomenon in a great number of adult languages, and even more universally attested in the language development of children. Newmeyer’s critique rightfully demonstrates the negative impact of the ad hoc construction of constraints, but his critique does not pose a serious challenge to the feasibility of ranking well-motivated and broadly attested syntactic and discoursal constraints with each other.

Third, one must question whether Newmeyer’s specific concerns about restrictiveness are as much of a problem for OT as they are for the Principles and Parameters framework that he assumes. This problem relates ultimately to the question of learnability, and though there is often an assumption that hypothesized grammars in language acquisition must be only those that are fully specified and admitted by UG, this is not a necessary conclusion. Tesar and Smolensky (2000), for example, point out that while such an assumption has the advantage of explaining why adult grammars invariably end up in the UG-allowed space, it has the potential disadvantage that it requires learners to be fully committed to a particular dimension of grammatical variation for which they have insufficient evidence. Such a position also does not deal as appropriately with adult L2 learner grammars, grammars that can, and often do, end up at variance with adult native speakers of the target language.
Finally, although in generativist circles it has become dogma that the job of linguistic theory is to restrict the space of grammars, this has popularly amounted to nothing more than delimiting a finite set of possible parameters with the associated assumption that the fewer possibilities, the better (i.e. more learnable) the theory. This popular notion is not informed by learnability theory. As Tesar and Smolensky (2000:2-3) have noted:

...limiting the set of possible grammars to a finite number serves only to improve the worst-case performance of the least informed learning method of all: exhaustive search, in which every possible hypothesis is examined...a grammatical theory with an infinite number of possible grammars might be well structured, permitting informed search that converges quickly to the correct grammar — even though uninformed, exhaustive search is infeasible.

Some adherents of PPT have not been adverse to invoking whatever new parameter seems to fit the needs of specific analyses, calling in question the ‘finiteness’ even of parameter theory, but as Tesar and Smolensky point out, even limiting a grammar to a finite number of parameters does very little to improve the learnability of a grammar, if the total possible number of grammars is quite high. Tesar and Smolensky deduce, ‘...a well-structured theory admitting an infinity of grammars could well be feasibly learnable, while a poorly structured theory admitting a finite, but very large, number of possible grammars might not’ (2000:2). Whether a grammar is ultimately learnable depends more on the learning mechanism than on the total number of candidates in a search space.

Learning research conducted in connection with principles and parameters (e.g. ‘cue learning,’ Dresher and Kaye 1990, or the ‘triggering learning algorithm,’ Gibson and Wexler 1994) has moved in one of two directions, both of which attempt to respect the autonomy of grammatical components: either they narrowly use the grammatical structure of a specific parametric system, or they use algorithms so general that they
apply to any (even nonlinguistic) parametric system. As Tesar and Smolenky (2000:4) note, such attempts come as a result of the fact that parametric systems have ‘little structure for the learner to exploit beyond the existence of a finite space for learning.’ In contrast, the OT approach advanced in this dissertation is a theory of crosslinguistic variation that uses a grammatically-informed learning algorithm that is demonstrably learnable.

Therefore, one may legitimately question whether the ‘problem’ of restrictiveness is a real issue or a pseudo-concern that arises from a particular set of generativist beliefs. It is true that OT allows a very large search space indeed. In fact, the number of potential grammars is equal to the number of possible total rankings \(N!\). But a process such as Constraint Demotion (Tesar and Smolensky 2000) drastically reduces the data complexity imposed by a large number of constraints, and this permits a useful type of restrictiveness not necessarily present in other generativist approaches (i.e. learners can efficiently converge upon target grammars through the use of the inherent structure provided by strict domination).\(^{18}\) Principles and parameters, despite its more ‘restricted’ learning space, has yet to demonstrate that it provides an approach to language acquisition that is actually learnable.

### 2.7.2 The functionality problem

Newmeyer’s second critique of OT is that some have attempted to use it to identify functional motivations for each of its universal constraints. Since Newmeyer rests firmly within the formalist camp, it is unsurprising that he calls this attempt ‘fundamentally wrong-headed’. Nevertheless, as noted earlier, researchers operating
within an OT orientation do span both sides of the functionalist/formalist divide. This is a natural outcome in that OT provides no fully elucidated theory of the universal constraints, but rather takes the constraints of particular grammatical theories and provides an organizing principle for them. For example, the SLA approach in this dissertation proceeds from a generativist view, assuming that candidate structures (constructed by GEN and considered by EVAL) conform to the basic architecture of the X-bar theory developed by Jackendoff in the 1970s and incorporated into GB theory in the 1980s. It differs from other generativist approaches, however, in that it shares with functionalism a rejection of the autonomy of syntax and a full appreciation for the interfaces that exist between grammatical components.

Nevertheless, there are those papers that view OT from a stronger functionalist orientation (e.g. Haspelmath 1999, Aissen 1999) and attempt to bridge formal and functional linguistics by finding constraint-function pairings that root OT constraints in their functional utility. Newmeyer’s critique here is not objectionable, and even he grants that functional motivations could play a role in language use and acquisition (with the assumption that questions of acquisition and use are outside the core of linguistic theorizing). Although considering why the constraints are the way they are may be an interesting diversion, these types of justifications are irrelevant if the theory assumes the innateness of constraints (as even Haspelmath 1999 admits). Functional motivations play little role in the work of the research program covered in this dissertation, and it is reasonable to concede the problem that Newmeyer raises regarding functionality, with the caveat that it applies only to some who work within OT syntax, and is not a failing of the theory itself.
2.7.3 Unmotivated competition

Newmeyer’s final criticism is his strongest attack on OT. Newmeyer rightly asserts that OT grammars fail to respect the autonomy of different grammatical components, a failure, he maintains, that obscures important generalizations about grammatical patterning. He labels this the problem of ‘unmotivated competition’ because it places grammatical components from differing levels in direct competition with one another. Newmeyer rightly observes that this competition is at the heart of OT, and is used to show that variances in hierarchical rankings account for differences between languages. However, he believes that this core notion leads to ‘ad hoc competition sets’, ‘uninsightful analyses’, and, even worse in his estimation, the importation of discourse constructs into syntax to handle discourse-dependent optionality, violating the received doctrine of syntactic autonomy. Newmeyer bases his argument on analyses of two types of optionality, discourse-independent, illustrated in 2.47 and discourse-dependent, illustrated in 2.48 (Newmeyer 2000:46).

(2.47.)
   a. I believe that it will rain tomorrow.
   b. I believe it will rain tomorrow.

(2.48.)
   a. You will never get me to eat fruit-flavored tofu.
   b. Fruit flavored tofu you will never get me to eat.

An OT grammar must be able to show how each of these types of optionality can arise. How it can accomplish this is not obvious, however, since OT grammars are deterministic, mapping each input into a single output that possesses the greatest harmony in relation to the ranking of constraints. If 2.47a and 2.47b are different in grammatical terms, and they are both generated from a single input, then the differences
leading to the inclusion or the exclusion of that must be due to constraints in the hierarchy.\textsuperscript{19} The choice of one or the other of these candidates appears to require that one is more optimal than the other, but every intuition of native speakers suggests they are in free variation.

OT researchers have addressed the issue of optionality in a variety of ways. Kager (1999:405) mentions that one terminological way of dealing with the issue is to redefine an output of the grammar as a set of forms rather than a single form, but concedes, ‘this does not solve the problem of how to generate sets of output forms’. Another approach, advocated by Legendre (2000) is based on the claim of Kroch (1989) that individuals have multiple grammars for their different stylistic registers. Legendre (2000) shows that in French \textit{wh}-constructions, level of formality determines the relative ranking of some constraints. Newmeyer opposes such an approach on the grounds that generative syntax has traditionally treated data from various registers of a grammar together, as with the \textit{wanna}-contraction or stranded prepositions. This is, of course, a very weak argument — that researchers should reject this approach because it is has traditionally been done another way.

True optionality can exist in OT if two constraints are crucially unranked in relation to each other. For example, assume that language $L$ has constraints $w, x, y$, and $z$. If the learner has received input that would require the demotion of say, $x$ below the other three constraints, and then further received input that would demote $y$ below the remaining two constraints, and then further received input that would demote $y$ below $x$, the constraint hierarchy would look as follows: $w, z \gg x \gg y$. This hierarchy leaves $w$ and $z$ unranked in relation to each other, leaving more than one possible choice as optimal in
this grammar. This type of free ranking permits an undetermination of the grammar. Kager (1999:407) notes that accounting for free variation may lead OT further in the direction of connectionism, where a numerical index indicating relative strengths in relation to other constraints may provide a probabilistic view of constraint interaction. This approach is similar to that of Anttila (1997), who also believes that variation materializes when the grammar underdetermines the output. As some researchers have argued, and I will later schematically illustrate, partial ordering of constraints is a possibility. For Anttila, preferences occur if these partial orderings are too weak to select a single winner, but still strong enough to leave a type of statistical fingerprint on the output. Preferred forms are then the result of one candidate beating another in a greater number of tableau, but still leaving the possibility for realization of the variant form.

Newmeyer dismisses true optionality as theoretically unlikely because ‘it is so easily sabotaged by the existence of some low-ranked constraint that distinguishes the two variants’ (2000:49), but a mere dismissal does not amount to evidence against it; furthermore, while it is true that in a totally ranked hierarchy (with no boundaries between components of grammars) such true optionality would be rare, the possibility for this type of optionality becomes far greater if the language learner does not have to consult the entire hierarchy for each piece of data, but is able to focus on the portion of the hierarchy relevant to a given structure. Exactly how this may be done within the learning algorithm is not fully clear, but the possibility of such a solution surely militates against viewing OT syntax as ‘inherently flawed’.

Numerous other possible solutions to optionality are debated among OT researchers, including ordered global constraint ties (2.49), ordered local constraint ties
(2.50), conjunctive local constraint ties (2.51), local disjunctive ties (2.52), and global disjunctive ties (2.53), here illustrated schematically (as adapted from Newmeyer 2000):

(2.49.) Ordered global constraint ties

![Diagram of ordered global constraint ties]

(2.50.) Ordered local constraint ties

![Diagram of ordered local constraint ties]

(2.51.) Conjunctive local constraint ties

![Diagram of conjunctive local constraint ties]

(2.52.) Local disjunctive ties:

![Diagram of local disjunctive ties]
The details of each of these theories are not critical here. The basic idea behind them is that crucial non-ranking of certain constraints in a grammar leaves room for some optionality. Newmeyer does nothing to refute these proposals himself; he relies on the debate between OT researchers on the advantages or disadvantages of certain approaches to suffice. His own position is simply that lower-ranked constraints must somehow destroy the possibility for an optimal choice, and that these types of approaches are difficult to reconcile with the learnability algorithms such as that of Tesar and Smolensky (2000). Once again, this type of argumentation does not approach any threshold of evidence for the invalidity of an OT approach, and it fails to understands how constraint demotion actually works. Constraint demotion does not function, as Tesar and Smolensky (2000:48) state, within ‘the confines of the space of totally ranked hierarchies’; rather, the algorithm operates within the larger space of stratified hierarchies.

‘Unmotivated competition’ between constraints is Newmeyer’s concern wherever the autonomy of syntax appears to be violated by OT. For optionality that does not involve differences in the information structure of the utterances, Baković’s (1997) analysis of optional complementizer deletion may serve as an example. Baković (1997:1) cites examples such as 2.54 and 2.55:
Sentences 2.54 and 2.55 show that complementizers in English are optionally present in some contexts. Baković argues that if the only constraints in a hierarchy were markedness constraints, such optionality would not be possible — the constraints would always compete, with the least-marked form invariably winning. However, markedness constraints may interact with faithfulness constraints, and under such conditions, optionality is a natural result. If we assume that specification for a subordination feature [SUB] is a part of the input that is received, then a ranking of faithfulness (ensuring that input and output values remain constant) over other markedness constraints would result in the following hierarchy, where optionality is the result of variance in the input, not a variable ranking of constraints within the grammar (Tableau 2.5):

<table>
<thead>
<tr>
<th>Tableau 2.5 Complementizer optionality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> [+SUB]</td>
</tr>
<tr>
<td>a. <em>I think that the coat does not fit him</em></td>
</tr>
<tr>
<td>b. <em>I think the coat does not fit him</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Input:</strong> [-SUB]</th>
<th><strong>Faith[SUB]</strong></th>
<th><strong>Markedness Constraints</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. <em>I think that the coat does not fit him</em></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b. <em>I think the coat does not fit him</em></td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

Newmeyer’s critique of this type of approach is stated but not argued. First, he objects to the expansion of the input to include functional contrasts. Second, he opposes the interaction between faithfulness and markedness constraints on the basis that it involves larger candidate sets and creates ‘vacuous ambiguity’ in well-formed sentences.
Neither of these objections strike the death-blow to OT that Newmeyer supposes. First, there is no reason to assume, a priori, that functional information cannot be part of a syntactic input. Indeed, a number of analyses within the mainline generativist approaches require a refined sensitivity to functional categories (Liceras 1988, 1989, Liceras and Díaz 1995, Pérez-Leroux and Glass 1997, 1999, Pérez-Leroux, et al. 1999), and if this sensitivity is not developed on the basis of input, it is unclear where it comes from. Second, the increase that faithfulness constraints create for the complexity of the candidate sets is only of concern if the exhaustive search method is used, but as Tesar and Smolensky (2000) have shown, the learning algorithm is significantly more sophisticated than that. It may be that Baković’s approach to handling discourse-independent optionality requires further development, but Newmeyer has not yet demonstrated this to be the case.

Newmeyer next considers ‘discourse-dependent’ or ‘discourse-sensitive’ optionality, optionality that occurs when the same input (which in OT syntax has been understood as lexical items together with their argument structure, tense, and aspect specifications) results in variable outputs on the basis of information structure. Newmeyer’s critique here is not unlike those already addressed. His primary concerns are that discoursal features are incorporated into the input, in the same way that information about functional categories was incorporated into the input in Baković (1997), and that the inclusion of such constraints expands the number of candidate sets beyond manageability.

For example, Newmeyer opposes Legendre’s analysis of French Stylistic Inversion (from Legendre 1999) on the grounds that allowing discourse or information structure constructs to compete with syntax poses a serious violation to the autonomy of
syntax. He argues that the principles characterizing well-formedness of sentences structurally must be different from those characterizing their appropriate use in discourse, using examples that place word stress on differing constituents, depending on the context, and arguing thereby that focus cannot be marked as a feature on a single constituent. He shows, for example that topics can be discontinuous (2.56), non-constituents (2.57), or not even need a syntactic representative (2.58):

(2.56.) Q: Are Mary and Tom friends?  
A: No, Mary HATES Tom.

(2.57.) Q: WHO did the man see?  
A: The man saw THE WOMAN.

(2.58.) Q: What’s the weather?  
A: It’s RAINING.

From this he concludes that topic is not a property that can be marked as a feature on syntactic constituents. Furthermore, Newmeyer argues that the relationship between syntax and information structure is indirect, with formal marking showing no need to correspond directly to information status. To illustrate, Newmeyer refers to Gundel's (1988: 214a) example of a formally definite NP that does not represent information known to the hearer (2.59):

(2.59.) I couldn’t sleep last night because the neighbor’s dog kept me awake.

Syntax and information structure usually have more than a one-to-one mapping. For example, Newmeyer cites Prince (1998) to show that processes such as left-dislocation or topicalization have multiple functions. Semantics can also diverge from both syntax and discourse, so in 2.60 two structurally different sentences
have the same information structure, although the same lexical item is focused in both sentences (from Valduví 1992:113):

(2.60.) a. *The boss hates BROCCOLI.*

b. *It is BROCCOLI that the boss hates.*

Given all of these considerations, Newmeyer draws the following conclusion (2000:60):

The built-in need of OT to regard each grammatical sentence as a winner in its own competition set has led to analyses that import aspects of information structure into the syntactic derivation. But the complexity and indirectness of the interface between syntactic structure and information structure shows that such a move is ill-advised.

On the face of it, Newmeyer’s critique appears very damaging to OT syntax and to the course pursued in this dissertation. The OT analyses used in this dissertation are guilty of precisely the critiques Newmeyer raises by grammatically encoding discoursal constraints and viewing them as universal elements of the OT grammar, and also by violating the autonomy of syntax by arguing that discourse and syntax interact as the second language learner moves towards acquisition of the target language. This is done despite the critique of Newmeyer for the following reasons. First, although it may be granted that discourse function is not invariably tied to a single constituent; for any given linear ordering of lexical items, there are clearly numerous focus possibilities, dependent on the discoursal requirements of the context. When a lexical item such as ‘broccoli’ in 2.62 is able to receive focus in more than one position in the sentence, this does not demonstrate that there is no interaction between components of grammar; rather, it shows that the significant interaction in this case is between information structure and stress assigned via the phonological constraints of that language. After all, not every significant
interaction is based on syntax. Rather than arguing for the autonomy of syntax by such
examples, Newmeyer is further demonstrating the need to examine interfaces between
grammatical components.

Second, Newmeyer’s claim that interactions between discourse and syntax are
complex and, at least apparently, indirect can also be granted. It is precisely this
complexity, together with the artificial requirement to separate syntax from other
components of grammar, that underlies the failure of the mainline generativists to be able
to account for various grammatical phenomena, of which pro-drop is a prime example.
The claim that allowing discourse to interact with syntax in the grammar makes the
grammar too complicated, or the constraints too numerous, is countered by the charge
that refusing to allow discourse to interact with syntax yields only an overly simplified
account that is non-predictive and must fall back on optional rules that ultimately yield
only description, not explanation. The presence of complex principles of interpretation
relating syntactic structure and information structure is unsurprising, nor should the
complexity of these principles dissuade researchers from making progress where
possible.

Third, the observation that syntax and information structure usually have more
than a one-to-one mapping does not argue against an OT approach. OT presents at least
the possibility of analyzing, in a way that is not possible within some other frameworks,
the host of constraint interactions that may demonstrate why syntax and information
structure do not have one-to-one mapping in certain cases. At best, mainline generativist
accounts yield no better explanations in these situations; at worst, they do not possess
even the means for addressing them.
Finally, while Newmeyer’s defense of the autonomy of syntax permits him to easily demonstrate that OT violates this generativist principle, his defense does not meet necessary criteria of explanation. The burden of proof for the principle of autonomy rests with those making the claim of autonomy. While Newmeyer alleges that this burden has been met, this assertion ignores data that do not fit the claim (such as that offered in the next section) and inadequately meets certain criteria of explanation. Contrary to Newmeyer’s claim, OT interactions between constraints can be precisely formulated and they make specific predictions, demonstrating a linkage between cause and effect. Furthermore, they result in measurable typological consequences.

Interestingly enough, Newmeyer’s critique wistfully longs for halcyon days when sentences displaying differences in free variation would be related via optional application of transformational rules. Even after the generative program in the 1970s led to a single movement rule, optionality could be handled by different applications of Move-α to a single D-structure, and interpretive principles applying at LF would leave the syntactic component untouched. Newmeyer views Minimalism, where all movement is obligatory, as a serious and unfortunate shift away from optionality. Apparent optionality in MP must now arise from two derivations from different candidate sets, a move which has its own unwelcome results. Therefore, according to him, both minimalism and OT share the problem of how to treat non-unique outputs. But returning to the optional application of rules is a dead-end street. Optional rules are neither predictive nor explanatory, and on this OT and Minimalism agree.
2.7.4 The syntax/discourse interface

The autonomy of syntax is an empirical issue; the fact that syntax often appears to make well-formedness judgments apart from other considerations does not establish autonomy. To be truly autonomous, it must *always* do so, and the failure of syntax to do this is can be demonstrated. Consider, for example, previous work in pro-drop, the focus of the research program discussed in this dissertation. If, as Newmeyer would have it, answers related to a syntax/discourse interface are not admissible as an explanation of the pro-drop phenomenon, some purely syntactic solution must be posited. Of course, this has been done. One standard proposal is that pro-drop and non-pro-drop languages differ in that non-pro-drop languages such as English have weak agreement features and do not permit the licensing of pro (Figure 2.5), but pro-drop languages such as Spanish have strong agreement features and do permit the licensing of pro at Spec of INFL (Figure 2.6):

Figure 2.5 Non-pro-drop languages: Weak agreement features
The problem with this type of analysis is two-fold: first, although this respects the autonomy of syntax, it is less descriptive and less explanatory than an account admitting discoursal factors. The explanation illustrated in Figures 2.5 and 2.6 does not account for other languages, such as Chinese, that have weak agreement and still have null subjects. Therefore, a disjoint solution becomes posited. In Chinese, so the suggestion goes (Huang 1984), pro is licensed by a discourse-bound operator in Spec of CP, and identified by null topics. So we now have three types of languages, those that license pro at Spec of CP, those that license pro at Spec of INFL, and those that do not permit the licensing of pro. Second, even this offered solution fails to specify where and when subjects will be dropped. It only describes how dropping, should it occur, would be licensed. This fails to be predictive, in contrast to OT accounts of the same phenomenon.
Grimshaw and Samek-Lodovici (1995) propose that subjects in a pro-drop language such as Italian are dropped when coreferent with the topic of the discourse. This clearly violates the principle of autonomy of syntax, but in so doing, it improves upon the previous analysis because it explains in terms of a discourse context precisely when a subject must be dropped and when a subject must not be dropped. It is, therefore, predictive and empirically testable. As Grimshaw and Samek-Lodovici maintain, null subjects are really null topics — an insightful claim that information structure requirements account for the missing subjects. In other words, where subjects are indexed with the discourse topic, they must be null. One reason for this may be that they are unambiguously recoverable, but regardless of the reason, the information structure requirements of the language requires the subjects to be missing, even when this creates an apparent violation of syntactic requirements (such as the unhedged version of the Extended Projection Principle).

Similar observations may also be made regarding focus. As we saw earlier, Grimshaw and Samek-Lodovici (1995) have shown that subjects which undergo what was previously believed to be free inversion are not free at all; these subjects are actually focused in the discourse. Thus, concerns for information structure relations places conditions on inversion. If focus, like topic, interfaces with syntax in significant ways to determine the acceptability of sentences, then the autonomy of syntax cannot be maintained. Recall that this was precisely the conclusion also of Costa (1997), who used Portuguese, English, Dutch, and Icelandic to illustrate differences between languages where there is a correlation between word-order and discourse functions and languages where this correlation is missing. Costa observed, and so should we, that syntactic
constraints are violated here only when they interact with discourse constraints. But the fact that these syntactic constraints are violated militates against any generativist approach in which such principles must be held inviolable. To account for apparent exceptions, other generative approaches are obligated to incorporate even greater optionality into the application of principles. Using the OT approach not only permits an account for the fact that syntactic principles are violated, but also permits us to identify the contexts where such violations occur, reducing the optionality present in the grammar.

For these reasons, OT provides an empirically cleaner, more elegant, and more predictive way of accounting for interactions between discourse and syntax. The theory is cleaner in that it does not need to feign inviolability of its grammatical principles, more elegant in that it can characterize differences in grammars as a simple reranking of the same universal constraints, and more predictive in that it envisages that the effects of a given constraint may be found independently of its being very high- or very low-ranked in a particular grammar and may specify just where we would expect to find such effects.

One final example of the interaction between discourse and syntax is that found in a recent study by Pérez-Leroux, et al. (1999) that shows how interfaces between discoursal and syntactic constraints in pro-drop languages control the distribution of null and overt pronouns. Pérez-Leroux, et al. (1999) show that in some Spanish cases syntax appears responsible for prohibiting the instantiation of a pronoun, and in other cases discoursal or contextual requirements determine the choice, as shown 2.61 and 2.62:

(2.61.) *María y Ø vamos de paseo.
   Maria and go-1pl on walk
   ‘Maria and I go for a walk.’
(2.62.) \textit{Ella/Ø no tiene dinero.}
\hspace{1cm} she-3sg not has money
\hspace{1cm} ‘She doesn’t have money.’

Pérez-Leroux, et al. (1999) observe that examples such as 2.61 demonstrate that coordination and null pronouns do not mix; whereas, in examples such as 2.62, it is the need for clarity that determines whether the pronoun may be null or overt. Examples such as 2.62 are additionally constrained by the fact that old information or discourse topics spark the use of null pronouns, whereas new or contrastive information normally requires an overt pronoun.

Pérez-Leroux, et al. (1999) do not appeal to OT to resolve this interaction between discourse and syntax; however, they represent a growing line of research in the last decade that has increasingly investigated interactions of discourse grammar with syntax in such a way that challenges traditional notions of the autonomy of syntax (e.g. Vallduví 1992, Erteschik-Shir 1993, Lambrecht 1994, Grimshaw and Samek-Lodovici 1995, Samek-Lodovici 1996, Costa 1997).

This section had the singular purpose of providing a defense of my research program, whose analytical approach uses OT to explain interactions involving a discourse/syntax interface, against the critique of those who challenge such an approach as a violation of the autonomy of syntax. Although the defense I have presented shows that Newmeyer has not persuasively demonstrated that OT syntax is inherently flawed, it should still obvious that there are essential issues remain to solved. This is unsurprising, since many of these issues have not been resolved (or even addressed) by other generative approaches, despite 40+ years of research within those paradigms. OT syntax is forging a new path to address these issues, and as with most intellectual pioneers,
researchers in this field also are encountering the disparagement of cynics. Despite this, the OT theoretical framework holds enough promise to warrant its continued exploration, even in the face of hearty criticism.

2.8 Limitations of the major proposals regarding pro-drop

In this section, I briefly summarize some of the limitations of the proposals that were reviewed in this chapter, and I suggest how future research may provide some remedy for these shortcomings. The first part of this chapter focused exclusively on pro-drop. Certain themes persist in the explanations that have been given. For example, from the earliest explanations of pro-drop through some of the more recent OT accounts, a great deal of attention has been paid to the importance of inflectional agreement or identification. Whatever the mechanism used for identification, most previous research relies upon the inflectional system and agreement markers of a language for answers, either by maintaining that the verbal inflectional system must be sufficiently rich, or by arguing that it must be sufficiently uniform.

Another recurrent theme in the literature, especially the early literature, is the optional application of rules to derive the effects of pro-drop. This line of research suffers from two limitations. First, it is questionable whether the earlier proposals that appealed to complex rule-based and transformational processes can adequately address the issue of learnability in acquisition. Later parameter-setting approaches appear to address this problem, but do so only by sacrificing empirical adequacy. For example, as noted earlier, although many researchers have assumed that null subjects and free inversion are triggered by the same parameter, there are certain Italian dialects that permit
inversion, but not null subjects, and there are languages such as Portuguese, that permit null subjects, but not inversion. To group disparate phenomena together under a single parameter, even when those phenomena are not necessarily linked in some languages and are acquired at different stages of a learner’s development in other languages, is no more explanatory than saying the learner once spoke a non-pro-drop language but now speaks a pro-drop language. Such a statement is non-arguable, but also non-explanatory. If it can be shown that the pro-drop phenomena are not a single parameter, but a fortuitous outcome of a particular grammar, then the validity of the parameter-setting approach to pro-drop is called into question.

Some of the earlier approaches held that clitics provide unique evidence that must be considered by a theory of null subjects; however, the line of reasoning followed by those who would invoke a ‘clitic-parameter’ should be scrutinized. Clitics do not behave in the same way as subject pronouns. For example, clitics can not be emphasized or coordinated, and they have a fixed order when they co-occur, all in contrast with pronouns. Also, they are by definition unstressed, whereas pronouns such as nosotros and usted(es) normally require stress. Furthermore, as Suñer and Lizardi (1995) note, pronouns, unlike clitics, can be separated from the verb by negation. Once again, applying specialized parameters to a problem does not actually solve anything; it simply stipulates an answer.

Previous research has certainly helped clarify some of the issues involved in pro-drop. Huang (1989) admitted discourse factors into the discussion of null subjects, an important contribution. But Huang (1989) does not go far enough. For example, he shows how discourse factors play a role in deciding why a null element cannot appear as
the subject of an English finite clause, but he says nothing about why Spanish finite clauses will sometimes permit null subjects and at other times prohibit them. Huang (1989) and most earlier accounts of pro-drop aim at the less ambitious task of predicting in which languages pro-drop can occur; but these accounts do not tell us when pro-drop will occur.

The Morphological Uniformity Hypothesis of Jaeggli and Safir (1989) also revealed an interesting correlation between some languages, but it does not say, for example, where or when Chinese or Spanish will choose to use null subjects. Furthermore, it does not explain why some languages with very uniform inflectional systems (e.g. German or Swedish) make little or no use of null subjects, and it fails to predict which languages will use null subjects. For example, Bani-Hassan Arabic does not uniformly mark verbal inflections, still it uses null subjects for some of its verbal forms but not others (Kenstowicz 1989). For this the MUH has no explanation. Studies by Lyons (1989), Platt (1989), and Davies (1996) provide direct empirical tests of, and counterevidence to, the MUH for adult second language acquisition.

The concerns of second language acquisition will be addressed in the next chapter, but here it may be said that if learner grammars in SLA adhere closely to UG principles, then these studies challenge Jaeggli and Safir’s hypothesis. A clear prediction of the MUH would be that no languages will surface that are both [–uniform] and [+null subjects]; however, this is not supported empirically. Lyons (1989) found no significant correlation between the admissibility of null subject in a second language and knowledge of the verb agreement system. Platt (1989) showed that even when learners recognized that English verbal paradigms were not uniform, they still dropped subjects in tensed
clauses. Davies (1996) found evidence that the recognition of non-uniform verb agreement morphology on the part of his learners was neither necessary nor sufficient to determine that null subjects should be banned in English. As Davies (1996) points out, among the available conclusions are either that UG does not play a role in SLA, or the MUH must be abandoned as part of UG.

As for the more recent accounts, Radford (1997) adds feature checking to the older stories of rich agreement and identification, but it remains unclear how feature checking will resolve the pro-drop question for languages with no inflectional system, and claiming that the need to check features can explain which languages will allow pro-drop does not yet explain why pro-drop will occur for a given utterance in any language. As for Radford’s (1997) example case, it should be noted that Old English had a number of properties that do make it difficult to classify it with standard pro-drop languages, undercutting part of the thrust of his argument. For example, inversion of pronominal subjects was greatly restricted in Old English (Allen, 1980), and null subjects were normally limited to expletives (Adams, 1987) and otherwise very restricted (Bresnan, 1976). These restrictions would not be predicted in the account of Radford (1997).

Speas (1997) begins well by looking at the interaction of violable constraints, but the constraints she chooses fail to account for interactions between syntax and discourse and then her account appears to fall back on the solutions (e.g. ‘rich agreement’) originally offered in other generative frameworks against which she is arguing. Nevertheless, Speas (1997) provides an example of how Optimality Theory could be used to approach this problem. Its fails only in that, while it does capture some generalizations
more elegantly, it does not produce a better explanation of differences between the grammatical systems of English and Spanish.

Grimshaw (1997) and Baković (1997) both contributed to a discussion of *that-trace* effects from an OT perspective, but these do not take it the additional step to show the predictions these types of constraints would make crosslinguistically. This problem is overcome in the approach of Grimshaw and Samek-Lodovic (1995), but although this study provides a theoretical analysis of differences between two language systems, it does not address developmental issues involved in the acquisition of the constraints responsible for null subjects, inversion, or *that-trace*.

It is a fair assessment of all the studies reviewed in this chapter to say that they make important contributions to linguistic theorizing concerning pro-drop, but they do not provide a developmental account of pro-drop in second language acquisition. For this reason, the next chapter reflects upon pro-drop research conducted within the field of second language acquisition.
Notes

1 Although most of the key terms in this chapter are defined, the discussion in this chapter assumes some familiarity with generative syntax. Those requiring a broader introduction to the goals and concepts of syntactic theory may wish to consult Cook and Newson (1996), Chapters 17-25 of Radford, et al. (1999), or the first six chapters of Fromkin, et al. (2000).

2 The concept of ‘parameters’ is discussed in greater detail in Section 2.3. Whether this parameter is referred to as ‘pro-drop’ (as was common in the early literature) or ‘null subject’ (as later became common), is of no significance here. Although the term pro-drop will be used throughout this dissertation, I will argue that the presence of an overarching pro-drop ‘parameter’ is something of a convenient idealization that researchers have used as shorthand to group languages that behave similarly. This dissertation will argue that there is no parametric setting that yields all of the properties shown above. Rather, these properties are acquired at differing times, in a specific order, and as a result of a predictable reordering of constraints within a grammar.

3 The concept of the ‘richness’ of a verbal paradigm has never been adequately defined. Presumably languages such as English and French (having only two or three inflectional distinctions for person and number) are not sufficiently ‘rich’, while languages that have fuller paradigms may pass some richness threshold.

4 It has been a common premise within syntactic theory that the human cognitive system interfaces with the articulatory-perceptual system (Phonetic Form, or PF) and the conceptual-intentional system (Logical Form, or LF). The elucidation of Chomsky (1981) given here is patterned after Rizzi’s discussion on the PRO’ Hypothesis (1982:159-161). It should be noted that this early account uses big PRO, the phonetically null element found in the subject position of control structures, as a label for the null subject. Chomsky (1982) later treated null subjects as a separate category, pro. Others continued to conflate PRO and pro into a single empty category, accounting for differences between them by invoking independent principles and parameters (e.g. Huang 1989, Borer 1989). More recently, Alexiadou and Anagnostopoulou (1998) argue that the EPP, and consequently pro, is no longer needed under Minimalism, and Petrovitz (1999) has argued against the existence of either PRO or pro categories altogether.

5 The details of each of these subtheories of GB (Chomsky 1982) will not receive detailed explanations here. ‘Case Theory’ broadly deals with the assignment of abstract Case, based on notions of ‘government’ found within GB theory. Under this theory, all non-empty NPs must be assigned Case. ‘Theta Theory’ deals with the assignment of thematic roles, requiring that arguments of verbs be assigned one (and only one) such role. ‘Binding Theory’ is concerned with the conditions that formally relate certain elements within a sentence. For example, as already mentioned, anaphors must be bound
and pronouns must be free in their governing categories, and referring expressions must be free.

6 Generative syntax has sometimes relied used an upside-down ‘T-model’ to illustrate the components of grammatical knowledge, as in Figure 2.7:

Figure 2.7  T-model of grammar

```
D-structure

| movement |

S-structure

PF component

‘sounds’

LF component

‘meanings’
```

In Figure 2.7, D-structure (underlying form) is related to S-structure (form of the sentence after movement; includes traces of moved elements), and S-structure is interpreted by the two components yielding phonetic and semantic representation.

7 Suñer preferred a PRO analysis because the presence of PRO allowed parallelism between missing subject elements for both matrix and embedded clauses. Suñer’s (1982) position still distinguished three types of null elements (1982:55-56, 74).

8 ‘Phi-features’ refers to the morphosyntactic indications of person, number, gender, and case.

9 Huang (1984) showed that the problem of identification here is restricted to null subjects. Although null subjects may be regarded as pro, Huang analyzes null objects as a variable A-bar bound by an empty operator. Condition C of Binding Theory would correctly predict it to be A-free.

10 In general, Safir (1985) argues that assignment of nominative case by a subject clitic to a postverbal NP permits inversion.

11 The term ‘features’ has played a central role in syntactic theorizing for some time. Grammatical features included person, number and gender (also called ‘phi-features’); the verbal features [± past][± tense]; and the binding features [± anaphoric] and [± pronominal] that stem from Chomsky (1981).

12 ‘Procrastinate’ is a principle favoring covert syntax over overt syntax, when there is a choice. The concept also involves the Minimalist idea of ‘Spell-Out’, which tells the grammar to move to create the Phonetic Form. Movement that takes place after Spell-Out is not as ‘costly’ as before Spell-out, because it does not have to pied-pipe phonological features (which are stripped away at Spell-Out), but delaying movement until after Spell-
Out can only take place with weak features. If strong features are involved, these are thought to be uninterpretable at PF, and therefore they must undergo deletion prior to Spell-Out.

13 This dissertation assumes that part of the universal construction of GEN is that it follows X-bar principles in generating outputs. This critically assumes that all languages obey X-bar principles, which is an empirical question for further research.

14 Grimshaw (1997:414) notes an exception for relative clauses, where ‘that’ is required to be present when it is the highest subject that is extracted.

15 A more detailed discussion of the problem of language acquisition is found in the next chapter.

16 Even the most ardent opponents of OT are often circumspect about attacking the value of OT for phonological research, perhaps due to the extraordinary and sweeping impact OT has had for that field.

17 I have used the term ‘formalist’ here rather than ‘generativist’ to not place OT outside of the generativist approaches. It is interesting to note that there are OT researchers in both of the camps contrasted here; for example, Haspelmath (1999) and Aissen (1999) use functionalist approaches within an OT framework, but other researchers such as Grimshaw (1997), Samek-Lodovici (1996), or Costa (1997) consider interactions between different levels of grammar without resorting to functional explanations.

18 Data complexity refers to the amount of data needed for an algorithm to ensure that it learns a correct grammar. Since the Constraint Demotion Algorithm (Tesar and Smolensky 2000) operates on informative data pairs, the maximal distance between an initial hierarchy and a target hierarchy is $N(N-1)/2$, where $N$ is the total number of constraints.

19 Exactly what constitutes input in OT has been a matter of some debate. This dissertation follows the view of Grimshaw and Samek-Lodovici (1995:590), who regard input as ‘a lexical head and a mapping of its argument structure into other lexical heads, plus a tense specification’.

20 There are contexts where the complementizer is obligatorily present or absent, and Baković accounts for this as well, but for the purposes of our discussion here, we will focus on the optionality issue.

21 Some researchers also include other elements (e.g. Baković’s [SUB]) as part of the input.

22 Legendre (1999) showed that in French, inversion indicates that the subject is ‘identifiable’ whereas non-inversion indicates that the subject represents active old information. Legendre thus uses information structure terms such as those discussed in
Lambrecht (1994) to construct a constraint that aligns non-active NPs with the right edge of $V^0$. This constraint, in competition with a constraint requiring subjects to be realized in Spec IP, yields the correct empirical results.

For example, left dislocation can be used to trigger a partially ordered set (poset), to aid in discourse processing or to permit island violations, and topicalization either triggers poset inferences or simply serves in the traditional topic-marking role.
Chapter 3

Theoretical Background II (Pro-drop in Acquisition)

3.0 Introduction

The goals of this chapter are (1) to introduce the major issues involved in acquisition research, particularly as these relate to the question of pro-drop, (2) to review relevant second language acquisition empirical studies that have been conducted on the pro-drop parameter, and (3) to review recent attempts to explain second language developmental issues related to pro-drop from an Optimality perspective.

Section 3.1 begins with a discussion of how second language acquisition studies have mirrored developments in linguistic theory more generally. This section will suggest that evolving developments in linguistic theory have been applied to specific, longstanding problems in linguistics, such as pro-drop, making these types of problems barometers of success for a theory. Section 3.2 considers the key acquisitional issue that any grammatical theory must address: the logical problem of language acquisition (Chomsky 1981). Section 3.3 then discusses two concerns related to the logical problem of language acquisition: the subset principle and learnability. The issue of learnability is central to any developmental account, and this dissertation is based on a particular learning algorithm emerging from research in learnability: the Robust Interpretative Parsing/Constraint Demotion Algorithm of Tesar and Smolensky (2000), for which this dissertation will offer further supporting evidence. Section 3.4 takes up two remaining critical assumptions for the argument developed in this dissertation: the accessibility of
Universal Grammar for L2 learners and the role of first language (L1) transfer in second language learning. Section 3.5 then reviews specific studies that have considered these issues through empirical tests of pro-drop from a parameter-setting perspective. These will underscore how the prevailing linguistic theory at the time of the research was empirically tested by developmental accounts. The need for a reconsideration of the pro-drop parameter from a developmental perspective, the focus of this dissertation, will become clear. Section 3.6 reviews the only SLA accounts related to pro-drop from an OT perspective to date: Park (2000) and LaFond, Hayes, and Bhatt (2001). As we will see, Park’s account was limited to null arguments in child second language acquisition, and LaFond, Hayes, and Bhatt dealt only with null subjects, one piece of the pro-drop phenomenon. Thus, this chapter sets the stage for remaining questions and a clear direction for further study. These questions and hypotheses, as well as the research design they suggest, are discussed in the next chapter.

3.1 Linguistic theory and SLA

This section situates the questions of SLA and the research agenda of this dissertation within the broader domain of linguistic theorizing. In the modern era, as Braidi (1999) and Mitchell and Myles (1998) have observed, L2-acquisition studies have mirrored developments in linguistic theory more generally. Early (pre-SLA) studies in the 1950s and 1960s were influenced by behavioral psychology and structural linguistics. The influence of Bloomfieldian views of language acquisition as habit formation were applied to language teaching, culminating in the development of the Audio-Lingual
Method, a method that involved drilling language patterns until proper language ‘habits’ were formed.¹

Researchers such as Fries (1945) and Lado (1957) claimed that by analyzing contrasts between native and target languages, problem areas for language learners might be predicted. Lado (1957:2) additionally claimed that learners ‘transfer’ the forms and meanings of their native language to the target language. Subsequent research supported claims for the existence of transfer, but overall, the ‘Contrastive-Analysis’ approach was shown to be less predictive than hoped.

Corder (1967) argued on the basis of Chomsky’s competence/performance distinction that learner errors provide a window into the learner’s linguistic knowledge. Like Chomsky, Corder sought to carefully distinguish unsystematic performance problems from the systematic errors that were a natural outcome of the system the learner was using. This approach treated errors not as the result of bad habits, but as clear markers of developmental stages of the language learner. Corder’s application of linguistic theory to questions of second language acquisition began a practice that has continued to the present day.

Both Contrastive-Analysis and Error Analysis shared a common assumption that learner grammars are deficient, transitional systems. This view was challenged by a third approach to L2 acquisition that viewed learner grammars as systematic grammars in their own right. Selinker (1972:214) describes these learner grammars as ‘a separate linguistic system based on the observable output which results from a learner’s attempted production of a TL [target language] norm’. Selinker brought the concept of ‘interlanguage’ to the fore of SLA studies and highlighted a number of processes that are
involved in the language acquisition puzzle: language transfer, overgeneralization, transfer of training, and L2 learning and communication strategies.

During the two decades following Selinker (1972), SLA researchers increasingly worked within the generative program and moved away from earlier ties to language pedagogy. Whereas earlier there had been an implicit assumption that SLA theories would provide teachers with recipes for successful practice, the radical shift away from practical requirements resulted in SLA developing as an autonomous field of inquiry with distinctive theoretical orientations, methodologies, and goals. As this field separated from language teaching, it began to explore various issues: To what extent does L2 learning mirror that of L1 learning? What role does the first language play in the acquisition of a second? Is access to Universal Grammar still available for the second language learner? What role do psychological, social, and environmental factors play in acquisition? To what degree are L2 learning and processing similar to other more general cognitive processes? Numerous questions have been raised; few definitive answers have been forthcoming.

Much of the early work in SLA attempted to apply research that had been done in L1 acquisition to an L2 context. For example, Brown’s (1973) morpheme study inspired investigations into developmental orders for the same grammatical morphemes among L2 learners. Dulay and Burt (1974) focused on child second language acquisition and found very similar acquisitional patterns in the L2 acquisition of English as for L1 acquisition. Bailey, Madden and Krashen (1974) extended this work to adult second language acquisition and found a great number of similarities to child L1 acquisition — those
acquiring English either as a first language or as a second language learn grammatical morphemes in a relatively set order, regardless of instruction.

Current work in SLA continues to explore important interfaces between L1 and L2 acquisition. For example, Brown’s (2000) research on Japanese, Korean, and Chinese learners of English as a second language suggests that in infant speech perception there is a direct link between the development of a feature geometry and the decline of perceptual capabilities. Brown posits that the acquisition of phonological structure imposes specific boundaries so that later input filters L2 allophones so that they are perceived in learners’ L1 phonemic categories. For example, Chinese speakers in this study differed from both Japanese and Korean speakers in their abilities to discriminate /l/ and /r/. According to Brown, this is expected because the presence of the [coronal] feature in Chinese permits a distinction on this dimension, while the absence of this feature in Korean and Japanese ‘funnels’ the acoustic signal for these two sounds into a single perceptual category (2000:40).

In another study, Young-Scholten and Archibald (2000), proceed from findings regarding the L1 acquisition of consonant clusters in certain quantity-sensitive Germanic languages to consider L2 interaction between segmental features and syllable structure. Young-Scholten and Archibald found that what is transferred from an L1 is not only the canonical CV structure, but also the complex interaction of the segmental inventory that determines the feature geometry of a segment, influencing what sequences of segments are allowed in the developing L2. Young-Scholten and Archibald also looked at sonority distancing and found that there is greater difficulty involved in acquiring L2 consonant clusters when contrasts in that L2 require an adjustment of the L1 feature inventory.
From the earliest SLA studies (Fries 1945, Lado 1957, Corder 1967, Brown 1973, Dulay and Burt 1974, etc.) to the present (Brown 2000, Young-Scholten and Archibald 2000, Park 2000, LaFond, Hayes, and Bhatt 2001, etc.), SLA studies have attempted to draw insights from L1 theorizing and, together with the specific insights that L2 research brings, draw conclusions about human processes of language acquisition. Each stage of the development of linguistic theory has seen a reinterpretation of earlier results in light of the current state of the theory, applied first to specific, persistent problems in linguistics (e.g. pro-drop), and then to the issue of second language acquisition.

Hence, the linguistic advance of a parameter-setting model (Chomsky 1980) came first, followed by various applications of this model to the pro-drop question (e.g. Rizzi 1982, Jaeggli 1982), and only later it was extended to questions of second language development (White 1985, Lakshmanan 1986, et al.). It was not until linguistic theorizing proposed that discourse issues must be admitted into an understanding of crosslinguistic differences (e.g. Vallduvi 1992, Erteschik-Shir 1993), that these concerns became incorporated into SLA research on pro-drop (e.g. Liceras and Diaz 1995, Pérez-Leroux and Glass 1997). The advent of Minimalism (Chomsky 1995) yielded another new theory of grammar that was applied first to pro-drop (Speas 1994, Radford 1997) and only later was followed by the predictable extension to an SLA context (Park 2000).

Given this pattern, it is natural to expect that the evolution of Optimality Theory (Prince and Smolensky 1993) should see the extension of this grammatical theory to the persistent question of pro-drop (e.g. Grimshaw and Samek-Lodovici 1995, Speas 1997) and then to the issue of pro-drop within SLA. But this final step of the progression has not yet been made for adult second language acquisition of syntax, with the exception of
LaFond, Hayes, and Bhatt (2001), and this paper addresses only the acquisition of null subjects in SLA. As is clear, pro-drop has been viewed by linguistic researchers as one of those persistent problems that a linguistic theory must be able to address, making it something of a barometer of the successfulness of a linguistic theory. This further confirms the need for a developmental account of pro-drop from an OT perspective.

Creating such an account of pro-drop may also address the need for an analysis that can explain important interfaces between discourse and syntax. Early SLA research followed the lead of the Chomskyan ‘revolution’ by considering how second language learners acquire parameters thought to be responsible for their competence. But as Chafe (1994) suggests, the initial Chomskyan approach was less ‘revolutionary’ than commonly imagined. According to Chomsky’s early formulation (1957:13), language is

...a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements...each language has a finite number of phonemes (or letters in its alphabet) and each sentence is representable as a finite sequence of these phonemes (or letters), though there are infinitely many sentences.

Although this approach was quite different than behaviorism, it still uncritically assumed that language was observable only via its forms, and that context specific meanings (i.e. pragmatics, discourse, social setting) were of little interest to any scientist attempting to explain core human competencies regarding language.

Recent developments in linguistic theory during the last decade suggest something far more revolutionary—that we need to recognize that language acquisition involves the complex functioning of a multifaceted, mental system. This means that a theory of grammar must, at least potentially, be able to both address interactions (and competitions) between various levels of language knowledge (e.g. conflicts between syntax and discourse) and to integrate questions of language learning and learnability into
a broader field of linguistic theorizing. Some theories of language (e.g. OT) are better prepared to meet this challenge than others.

3.2 The logical problem of language acquisition

Chomsky (1986) put forward a set of questions that he believed formed the essential framework for further inquiry into the phenomenon of language. They are questions that still are critical to linguistic theory and language acquisition today:

1. What is the system of knowledge?
2. How does this system of knowledge arise in the mind/brain?
3. How is this knowledge put to use in speech (or writing)?
4. What physical mechanisms serve as the material basis for this system of knowledge and this use of language?

Chomsky’s second question relates to what has been labeled the ‘logical problem of language acquisition’ (Hornstein and Lightfoot 1981). A clear formulation of this problem is found in Crain and Thornton (1998:283):

Many aspects of grammatical knowledge are represented as constraints, that is as sanctions against linguistic analysis of one kind or another. Constraints are negative statements. It is safe to assume that not all children, perhaps no children, encounter evidence pertaining to constraints. The pertinent evidence would be information about which linguistic expressions and meanings are prohibited in the target language. It follows from the absence of such negative evidence in children’s experience that knowledge encoded by constraints is not learned from experience. If not, then this aspect of linguistic competence must be innately specified, as part of Universal Grammar.

This problem is also referred to as the ‘poverty-of-stimulus’ argument or ‘Plato’s Problem,’ referring to a passage in The Meno where Plato records how Socrates led an uneducated child to discover theorems of geometry by only asking questions. Plato pondered how the child found the truths without being given any information.
concluding that this was evidence of an earlier existence, knowledge in the child’s mind that was reawakened when stimulated by questions. Chomsky argues (as Leibniz did) that Plato’s conclusion was essentially correct if by ‘preexistence’ we understand instead that certain aspects of our knowledge are innate, part of our genetic/biological endowment.³

The theorems of geometry are really quite basic compared to what children must learn about language. While children must be exposed to input before language acquisition takes place, once input is present, they acquire very complex grammatical principles (structure dependency, subjacency, binding, etc.) in a brief period of time, without explicit instruction and on the basis of input that is impoverished, imprecise, and variable. For example, Lakshmanan (1994:3) shows that children consistently choose 3.1b, which applies a computationally complex structure-dependent rule, over 3.1c, which uses a computationally simple structure-independent rule.

(3.1.) a. The book which is on the table is dull.

b. Is the book which is on the table dull?

c. *Is the book which on the table is dull?

Somehow, children know that what moves to the front of interrogatives of this type is the verb in the matrix clause, rather than simply moving the first verb in a linear ordering. Children also, without instruction, learn that ‘them’ in 3.2a can be coreferential with ‘the men’ in 3.2a, and also that it could refer to some other contextually salient persons, but that in 3.2b ‘them’ permits only the latter interpretation.

(3.2.) a. I wonder who the men, expected to see them,
b. *The men expected to see them.

How do children know these fine points related to the syntactic ordering of elements, the interpretation of pronouns, and semantics?

Fromkin (2000:472), notes that, in Italian, participle constructions can contain both atelic verbs, as in 3.3a and telic verbs, as in 3.3b.4

\[(3.3.)\] 
a. *Gianni ha ballato (per un ora)  
‘John has danced (for an hour).’

b. *Gianni è caduto (*per un ora)  
‘John fell (*for an hour).’

Without explicit instruction, 2-year-old Italian children restrict their use of the past participle construction to telic predicates such as *cadere ‘fall’, *rompere ‘break’, *arrivare, ‘arrive’. They do not use this past participle construction with atelic predicates such as *volare ‘fly’, and *ballare ‘dance’. How did Italian children acquire the telic/atelic distinction, and why do they undergeneralize at this stage of their acquisition?

In the area of morphology, ever since Brown’s (1973) study on the order of acquisition for 14 grammatical morphemes it has been repeatedly demonstrated that children’s acquisition follows clear developmental patterns. Mitchell and Myles (1998:27) point out that, regardless of the language being learned, children begin to use negation around the same age and mark the negative in similar ways, by first attaching some marker to the outside of the sentence and only later moving this marker inside the sentence. Since adult grammars differ cross-linguistically in the way negation is handled, why do children learning various languages follow similar developmental paths?
Given all these types of evidence from various areas of children’s grammars, we are left with the question of how children come to know these subtle distinctions and nuances of language without being taught. Answers pointing to mimicry of parents, the hearing of simplified input, or explicit correction are all easily repudiated. Arguing that children do not produce certain agrammatical forms because they have not heard them before does not explain why children will produce forms such as ‘Joey goed there with me.’ The types of developmental errors children make on their path to adult language acquisition are predictable, and are little affected by any overt attempts to instruct them.

This, then, is the logical problem of language acquisition, a problem that generative frameworks specifically attempt to address. For Chomsky, the solution to Plato’s problem is simply that the properties of the mind/brain include certain principles of the language faculty — UG (or a Language Acquisition Device, a linguistic ‘black box’) — that, given a rich enough linguistic environment, is able to determine the value of certain parameters and provide interpretation of linguistic expressions, even those that a child learning a given language has never encountered.

Cook (1991) outlines four stages to the poverty-of-the-stimulus argument. Given a particular aspect of linguistics knowledge, for example pro-drop, the argument would be stated as follows (adapted from Cook 1991:83-84):

Step A: Native speakers of a language have specific intuitions about certain aspects of the pro-drop.

Step B: These aspects of grammar could not have been acquired from the language input typically available to children.

Step C: These aspects of grammar have not been taught.

Step D: Therefore, these aspects must be built into the mind.
As the stages show, linguistic theory begins by looking at the knowledge of the native speaker. Since the linguistic input the child receives is insufficient to teach this principle, other sources must be sought. Cook and Newson write, ‘The poverty-of-stimulus argument is fundamentally simple; whenever you find something that the adult knows which the child cannot acquire, it must already be within the child’ (1996:85).

Although the logical problem of language acquisition initially arose from a first language perspective, SLA researchers have argued that it applies to second language contexts as well (White 1985, 1989, Flynn 1987, Cook 1996). As in L1 acquisition, L2 learners acquire the grammatical properties of the target language in the face of insufficiently rich and insufficiently precise input. For example, Spanish learners of English as a second language must learn that English is not a pro-drop language, despite the fact that the input they receive in informal conversation, email, etc., often uses subjectless sentences.5

The logical problem of language acquisition has led second language researchers to ask specific questions related to the relationship between L1 and L2 acquisition. They have queried to what extent the L2 learner has access to UG, how the L1 interacts with the acquisition of an L2, and how we might account for the variation in ultimate attainment of a second language. For Chomsky, the differences between L1 and L2 acquisition are greater than the similarities. It is true that while first language acquisition happens naturally and with little apparent effort, learning a second language for adults requires much more effort with uncertain results. Nevertheless, for most of the world, learning second languages is a normal human activity, and Plato’s problem is no less
important in second language contexts than for first language. We must somehow account for how a second language learner acquires constructions that could not otherwise be reasonably acquired on the basis of the input alone. This implicates some role for UG also in second language learning. As Flynn and O’Neill (1988:2) claim, SLA legitimately interfaces with linguistic theory in its ‘attempt to examine the extent to which a theory of UG could be useful in explaining the L2 acquisition process...’ and in its desire to identify ‘evidence for the role of UG in L2 acquisition’ and the domains in which this evidence may be found.

3.3 The Subset Principle and learnability

Since learning a grammar is difficult (at least for adult second language learners), and the input available to language learners is not sufficient to achieve this learning, the logical problem of language acquisition raises two additional important issues: the subset principle and the issue of learnability. These issues involve questions concerning the restriction of the learning space of possible grammars and the search method learners must use to arrive at the target grammar.

One attempt at a comprehensive theory of learnability for L1 acquisition is found in Pinker (1996), who argues that children acquire language ‘by exploiting rich formal and substantive constraints on the types of rules that languages may have’ (1996:358). Pinker claims that children entertain a small subset of the possible hypotheses consistent with the input they receive, with certain triggering conditions sparking reevaluation of the hypotheses. This reference to the Subset Principle (Angluin 1978, Berwick 1985, Manzini and Wexler 1987) is important also for L2 acquisition.
The subset-superset relationship is one that has played an important role in parameter-setting accounts, which hold that children choose the parameter with the fewest possible assumptions, based on the language input received, shown in Figure 3.1.

Figure 3.1 Subset relationship

In Figure 3.1, ‘Language A’ is in a subset relationship to ‘Language B’. This means that all possible sentences in ‘Language A’ are also possible in ‘Language B’, but the reverse is not true. The assumption, then, is that children must begin with the most restrictive grammar (enlarging it as the data permits), because if they were to start at the larger grammar, they would never receive enough evidence to restrict it to Language A, since the sentences found in the smaller grammar are also found in the larger grammar.

The subset principle has led some researchers (Berwick 1985, Phinney 1987, et al.) to assume that non-pro-drop languages are subsets of pro-drop languages. This permits predictions to be made regarding the difficulties involved in taking a particular learning path. For example, speakers of English (the subset language) would only require positive Spanish input (e.g. a sentence with a null subject) to begin to realize that they must adopt a larger sphere of language possibilities than their L1 requires. However,
learners of Spanish are in a more difficult situation. Positive input from English will yield overt subject pronouns, but these pronouns are also allowed in Spanish. Spanish speakers, being in a superset relationship to English, will require negative evidence that their pro-drop sentences are not grammatical in English. The requirement for additional negative evidence should make the learning process more difficult.

While some researchers believe the subset principle applies to the pro-drop relationship between Spanish and English, others have challenged this on several grounds (e.g. Hyams 1986, Wexler and Manzini 1987, MacLaughlin 1995). First, the subset condition is not fully met. For example, both English and Spanish use overt lexical subjects, but only English uses expletive subjects and only Spanish uses null referential subjects. This leads to an intersecting relationship between two languages such as Spanish and English, illustrated in Figure 3.2.

Figure 3.2 Intersecting relationship of subjects in Spanish and English

Second, even the expanded diagram in Figure 3.2 does not take into account that learning Spanish involves more than a general acceptance of null subjects. As Galván (1998) points out, learners of Spanish must learn both that null subjects are possible, and that certain discourse factors constrain their use. Any inclusion of discourse conditions would require Figure 3.2 to be amended so that nontopic null subjects would fall in the
intersection between the two languages, but null topic subjects would fall only on the Spanish side (Figure 3.3):

**Figure 3.3 Discourse conditions in Spanish and English**

![Venn diagram showing null topic subjects, overt non-topic subjects, and expletive subjects](image)

Finally, although the Subset Principle provides one way to restrict the hypothesis space for learner grammars, it does not necessarily hold that this greater restrictiveness will automatically translate into a more learnable grammar. For learnability purposes, the manner of search is more important than the total size of the search area. Consequently, the choice of search method becomes critical in a developmental account of L2 learning.

The learning algorithm used in this dissertation, that of Tesar and Smolensky (2000), is able to arrive at target grammars without reference to the Subset Principle. It does so even assuming no variance in the set of possible inputs to the grammars of all languages; when the grammar of a given language is supplied with this set of universal inputs, the grammatical inventory of that language is defined as the output forms that emerge as a result of the operation of the grammar. This concept is called the ‘richness of the base’ (Prince and Smolensky 1993).

Although the learning algorithm of Tesar and Smolensky is not obligated to assume a particular initial state (i.e. it arrives at the target hierarchy regardless of the initial hierarchy assumed), an initial state could be hypothesized. Tesar and Smolensky
report on a suggestion made to them in personal communication with Prince (1993) that the initial hierarchy of L1 learners may be such that faithfulness constraints are lower ranked than markedness constraints, with the result that structural constraints are demoted below faithfulness constraints only ‘in response to the appearance of marked forms in observed overt structures’ (2000:76).

Tesar and Smolensky (2000) argue, however, that it is not the Subset Principle that is critical to the learnability of a grammar, but rather the particular learning mechanism that is employed. They focus on a particular problem in language learning — how learners, who often receive overtly ambiguous language data, are faced with a serious paradox: they cannot determine a grammar’s hidden structure until they have constructed a grammar based upon their interpretation of the overt forms they hear, but they cannot construct a grammar without some analysis of the hidden structure. To address this paradox, Tesar and Smolensky have proposed a learning procedure where learners’ first guesses at a structural analysis are used to improve their grammar, and this improved grammar is then used to improve the analysis. In other words, through successive approximation, learners acquire progressively better interpretations and a progressively better grammar simultaneously.

Tesar and Smolensky (2000) look to OT for the core principles that inform this learning strategy, and in their proposed model, Robust Interpretive Parsing / Constraint Demotion Algorithm (RIP/CDA), and they provide evidence for the accuracy and computational efficiency of their proposed model through a series of computer simulations and by a set of formal proofs. Their central claim is that OT provides the learning mechanism (RIP/CDA) through which the interdependence of grammars and
structural descriptions is overcome, allowing the learner both to assign structure and to learn grammar at the same time. The learning problem is decomposed into several parts—deducing hidden structure in language data, using the data to improve the existing model, assigning an improved hidden structure to the original overt data, and once again learning the grammar (using a ‘robust’ parser). This divides the problem into one of parsing and grammar learning.

The centerpiece of Tesar and Smolensky’s theory is ‘Constraint Demotion,’ the notion that constraints violated by grammatical structural descriptions must be demoted, in the total ranking of constraints, below constraints violated by competing (ungrammatical) structural descriptions. The CDA operates in the following manner: When the hearer attempts to create a grammar based on the output of the speaker, Robust Interpretative Parsing first computes an input for the speakers’ productions. The hearer’s grammar then compares constraint violations of the target optimal output (winner) to those of the current grammatical system (loser). The algorithm cancels out the constraints violated by both winner and loser candidates as illustrated in 3.4:

(3.4.) \[\text{<input A>}\]

| Candidate X (loser): Constraint 2, Constraint 3 |
| Candidate Y (winner): Constraint 1, Constraint 3 |
| Candidate X (loser): Constraint 2, Constraint 3 |
| Candidate Y (winner): Constraint 1, Constraint 3 |

The hearer’s grammar then demotes the constraints violated by only the winner below those violated by the loser, resulting in a grammar that will produce the intended outputs with fewer violations. This process of constraint demotion proceeds recursively until there are no more mismatches between the perceived output and the grammatical system.
The algorithm that Tesar and Smolensky propose accomplishes more than traditional parameter-setting models. It reveals not only a workable way to arrive at a desired end state, but also in that it provides a clear description of the developmental steps expected as individual constraints are demoted in the total hierarchy. But does parameter setting provide a more learnable model, since it is usually thought to contain only binary choices for a limited number of parameters? Does not the OT model require that more interactions are involved, resulting in a far larger number of possible grammars and, thus, a less learnable system?

Tesar and Smolensky answer these concerns by demonstrating that, although the total number of possible rankings in an OT system may be quite high with even a limited number of constraints, the restrictiveness of the structure OT places on the grammar permits learners to efficiently arrive at a target grammar in a reasonable number of learning steps. Parameter-setting models, even with their more limited set of possible grammars, do not improve upon the learnability of a language, if an uninformed learning method is used — an exhaustive search of all possibilities:

Comfort from the finiteness of the space of possible grammars is tenuous indeed. For a grammatical theory...might be well structured, permitting informed search that converges quickly to the correct grammar—even though uninformed exhaustive search is infeasible...a well-structured theory admitting an infinity of grammars could well be feasibly learnable, while a poorly constructed theory admitting a finite but very large number of possible grammars might not.

(Tesar and Smolensky 2000:2-3)

Tesar and Smolensky remark that a parameter-setting approach with \( n \) parameters admits at least \( 2^n \) grammars (assuming only binary parameters). This means that the search space increases exponentially with each proposed parameter, and the exhaustive search method quickly becomes unfeasible.
Tesar and Smolensky are not the first to observe this problem, and there are other proposals to deal with issues of learnability within a parameter-setting framework. For example, Dresher and Kaye (1990) proposed a ‘cue learning’ approach that could be applied specifically to learning metrical stress. But the use of a particular algorithm for each component of grammar is less preferable than an approach that can account for multiple components of a linguistic theory.

Other proposals, such as the Triggering Learning Algorithm (Gibson and Wexler 1994) or modifications of it (e.g. Niyogi and Berwick 1996), may be applied to a more general class of components, but as Tesar and Smolensky note, these algorithms are minimally informed by grammatical theory. Gibson and Wexler’s algorithm randomly flips parameters to arrive at an analyzable input, and in Niyogi and Berwick’s version, the randomly flipped parameters do not even need to directly result in analyzability. Consequently, these approaches are simply generic search algorithms that could be employed for any parameterized system; they make no use of grammatical theory or its unique properties. In contrast, learning in Tesar and Smolensky’s algorithm is derived solely from general grammatical structure and informed by a specific theory of grammar. The particular strengths of their algorithm make it the choice for this dissertation’s analysis of Spanish L2 data.

3.4 Theory of access to UG and L1 transfer

The analysis provided in this dissertation will critically assume both that learners of an L2 have access to Universal Grammar (through the OT learning mechanism), and that the initial hierarchical ranking of constraints for these learners is the ranking of their
L1 (i.e. ‘transfer’). Neither of these positions has unanimous support among SLA researchers; therefore, this section briefly discusses these ideas.

Although generative linguistic research has proceeded from the assumption that humans have innate access to Universal Grammar (Chomsky 1965, 1975, 1986) for their acquisition of their first languages, access to UG for second language acquisition has been more controversial, and differing positions have been taken as to the extent to which learners have access to and use UG as they learn a second language. In terms of access, three broad positions emerge: first, the position that, unlike learning of the L1, no access to UG is available for adult L2 learners (Lenneberg 1967, Clahsen and Muysken 1986, Bley-Vroman 1989); second, the position that learners have access to UG, but only partially, not in the same direct and unmediated way that there is access to UG for the acquisition of the first language (Schachter 1989, Strozer 1992, Vainikka and Young-Scholten 1991, Bhatt and Hancin-Bhatt 1997); finally, the position that L2 learners have full and direct access to UG, making this one way in which L1 and L2 acquisition are similar (Epstein, Flynn and Martohardjono 1996, Lakshmanan 1993, Schwartz and Sprouse 1996).

Section 3.1 stated that the early Contrastive Analysis and Error-Analysis theories were challenged by the interlanguage approach of Selinker (1972), who viewed learner grammars in terms of a complex interaction between first language influences (transfer) and innate language learning processes (access to UG). This did not involve a complete disregard of first language influences, and, in fact, Gass and Selinker argued for such influences. They claimed that the ‘overwhelming evidence that language transfer is indeed a real and central phenomenon that must be considered in any full account of the
second language acquisition process’, and they further argued that SLA involves hypothesis testing using UG, L1 data, L2 data, and the knowledge of interlanguages (1983:7).

Nevertheless, debate concerning the notion of transfer continues. For example, Martohardjono & Flynn (1995) analyzed control structures in English, Japanese, Chinese and Spanish and concluded with the strong claim that L2 learners ignore their L1 syntax and rely instead solely upon principles of UG. In contrast, White (1989) argues that despite the learner’s use of UG, the L1 does indeed play a significant role in L2 acquisition. White claims that L1 parameter settings are part of the second language learners' interlanguage, and that they influence attempts to both understand and produce the target language.

White (1989:48-9) outlined five logical possibilities for the relationship between transfer of the L1 and the accessibility of UG for L2 acquisition (Figure 3.4):

Figure 3.4 Possible relationships between access to UG and L1 transfer

a. UG is accessible and functions as it does in L1 acquisition.

b. UG is accessible, but learners initially transfer the settings of the L1.

c. UG is accessible, but only via the settings of the L1.

d. UG is accessible, but does not function identically as in L1 acquisition.

e. UG is inaccessible.

These possibilities are exemplified in Table 3.1 for a parameter-setting model of pro-drop where the pro-drop parameter is given two settings [+pd] or [-pd] (Adapted from Braidi 1999:62-64):
Table 3.1 Parameter model of acquisition of pro-drop

<table>
<thead>
<tr>
<th>Condition: UG is accessible and functions as it does in L1 acquisition.</th>
<th>L2 French Input [-pd]</th>
<th>→</th>
<th>L1 Spanish UG Default [+pd]</th>
<th>→</th>
<th>Set to [-pd]</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Condition: UG is accessible, but learners initially transfer the settings of the L1.</th>
<th>L2 French Input [-pd]</th>
<th>→</th>
<th>L1 Spanish [+pd] UG</th>
<th>L1 transfer [+pd]</th>
<th>Reset parameter to [-pd]</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Condition: UG is accessible, but only via the settings of the L1.</th>
<th>L2 French Input [-pd]</th>
<th>→</th>
<th>L1 Spanish [+pd] UG</th>
<th>L1 transfer [+pd]</th>
<th>Resetting unnecessary</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Condition: UG is accessible, but does not function identically as in L1 acquisition.</th>
<th>L2 French Input [-pd]</th>
<th>→</th>
<th>UG &amp; other components</th>
<th>→</th>
<th>Variable success in acquisition of pro-drop due to other components</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Condition: UG is inaccessible.</th>
<th>L2 French Input [-pd]</th>
<th>→</th>
<th>L1 Spanish &amp; general learning mechanisms</th>
<th>→</th>
<th>Variable success due to reliance on general learning mechanisms</th>
</tr>
</thead>
</table>

Table 3.1 assumes a default setting of [+pro-drop] for a pro-drop parameter. How the learner grammar will handle L2 input depends both on the transfer/access possibilities and on the relationship between the default setting of the parameter, the setting of the parameter in the L1, and the setting of the parameter of the L2. The parameter may require ‘setting’, ‘resetting’, ‘maintaining’, ‘transferring’, or some other treatment.

White (1989) harmonizes with Schwartz and Sprouse (1996) who hold, in contrast to Martohardjono and Flynn (1995), that all L1 parameters initially transfer and that it is only due to the failure of the L1 grammar to adequately represent the facts of the L2 that learners must restructure their grammar with the options available to them from UG.
This position, referred to as Full Transfer/Full Access, was supported in Schwartz and Sprouse’s study by spontaneous production data of an adult native speaker of Turkish who was learning German. Turkish and German are very dissimilar in surface syntax, and as this position would predict, the learner was quickly forced to UG-constrained restructuring of his syntactic system. Selinker and Lakshmanan (1994) provide numerous additional examples from Dutch, Czech, Spanish, French, Hebrew, and Hindi, to show that transfer of L1 syntax to the L2 not only occurs, but also prolongs the restructuring, particularly when there are multiple effects at work.

Although questions related to transfer and access have regularly been posed within the parameter-setting framework, the same questions are applicable to the OT approach used in this study. From an OT perspective, although there are no parameters to transfer, what may transfer are the constraint rankings of the L1. In terms of access to UG, from an OT perspective this could mean either access to default constraint rankings supplied by UG or, more broadly, to the UG instantiated learning mechanism, represented by the operation of the OT grammar. Since there is no a priori reason to assume a default ranking of constraints, and since the successful operation of the RIP/CDA learning algorithm does not depend on an initial default ranking, the analysis provided in this dissertation will assume that while learners have full access to UG, UG is defined as the OT grammar learning mechanism together with the universal set of constraints.
3.5 Pro-drop in SLA

This section reviews the findings and methodologies of several SLA studies testing aspects of the pro-drop. For most of the early studies (e.g. White 1985, Emberson 1986, Hilles 1986, Phinney 1987, and Liceras 1988, 1989), this involved establishing the facts of pro-drop, submitting theoretical parametric accounts to empirical testing, and deciding on the default option for the proposed ‘pro-drop parameter.’ Very few SLA studies related to pro-drop have moved beyond the parameter-setting model; these few will be covered in Section 3.6. This section concludes with a short statement of the set of facts that SLA research on pro-drop has revealed, as well as remaining questions these studies leave unanswered.

The concerns of several early SLA researchers revolved around extending the L1 acquisitional work of Hyams (1983, 1986) to an L2 context, with some of the Hyams’ hypotheses either carried into, or challenged in, these L2 studies. For example, whereas Hyams (1983) argued that pro-drop was the unmarked setting and that, if the target language was not pro-drop, a switch of parameter settings was required, White (1985) argued that the unmarked setting is [-pro-drop].

White (1985) argued that learners of a language at variance with their L1 in regards to pro-drop do not immediately or easily reverse this parameter when they begin to acquire the L2; rather, they initially transfer the setting of their L1 into the L2. White was also concerned about the types of input needed to trigger the change.

White’s study tested the acquisition of English, specifically related to the three characteristics most closely associated with pro-drop: null subjects, subject inversion, and that-trace effects. Her study involved 73 adult L2 learners of English at McGill
University in Montreal, Canada. Of these 73 learners, 54 were native speakers of Spanish and 19 were native speakers of French. White used a grammaticality judgment task where subjects were given 31 English sentences, some of which were well-formed English sentences and others that were not. The ill-formed sentences contained null subjects, inverted subjects, and violations of that-trace. White hypothesized that Spanish speakers would be likely to reject an English sentence such as (3.5) because the Spanish equivalent in (3.6) requires the presence of the complementizer que.

(3.5.) Who do you believe will be the next president?
(3.6.) ¿Quién crees *(que) será el próximo presidente?

In contrast, White thought that French speakers would be less likely to make pro-drop related errors in their assessment of English sentences, because French and English share the characteristic of both being non-pro-drop languages.

White’s results showed that Spanish speakers did indeed have more difficulty than the French speakers in correctly judging grammaticality when null subjects were used, although both groups also had difficulties judging sentences with expletive subjects (‘It seems that Fred is unhappy’). Both the French and Spanish groups performed well in regards to subject inversion and poorly in regards to that-trace. Although White concludes that pro-drop is a parameter with a set of related consequences, the results of her study clearly show that the various pieces of pro-drop are not simultaneously acquired.

Emberson (1986) argued against access to UG in second language acquisition. Emberson focused on acquisitional differences between ‘core’ (i.e. pro-drop) and ‘non-
core’ (i.e. the present progressive tense) aspects of Spanish. The study involved 41 English speakers in a fourth-semester Spanish class at the University of Texas in Austin. Emberson reports on correlations of student scores on eight grammaticality judgment tests.

Emberson concluded that UG is not active in second language acquisition, based both on a lack of correlation between aspects of pro-drop and on better performance on the progressive tense (a non-core aspect of the grammar) than on pro-drop (a core aspect of Spanish grammar). This, of course, critically assumes that a core/non-core distinction exists for the Spanish progressive and pro-drop, an assumption not shared elsewhere in the pro-drop literature. Emberson’s results do, however, indicate an important place for transfer, since L1 English learners of Spanish did better on the present progressive (a feature that Spanish and English share) than they did on pro-drop related tasks (where differences exist between Spanish and English). Rather than demonstrate that UG is no longer active in second language acquisition, Emberson simply provides more evidence of a legitimate role for transfer in SLA.

The study of Hilles (1986) offered another test of Hyams (1983) regarding the triggers involved in resetting the pro-drop parameter to [-pro-drop]. Hyams had argued that the presence of expletives reveal to the learner the need for parameter resetting. Hilles (1986) pursues the same line of reasoning as Hyams (1983) but departs from the earlier hypothesis by posing that the presence of modals, rather than expletives, supply the triggering data. Hilles attempted to determine when lexical subject pronouns first surface through the use of a longitudinal study of a single subject, Jorge. Jorge was a 12-year old native speaker of Spanish who had not formally studied English. From a
mixture of elicited and spontaneous data collection methods (none rigorously defined), Hilles gathered information regarding the number of instances of missing subjects. She also attempted to count the instances where pro-drop could have occurred (if the text had been in Spanish) but did not occur in English. Hilles then developed a formula by which she divided the total number of null subject instantiations by the sum of that number added to the number of instances where null subjects would have been expected to occur in Spanish. Interpreting the data through the use of this formula, Hilles concluded that Jorge’s use of null subjects was declining over time.

The remaining question for Hilles concerned what triggered the parameter change. Hilles noticed that the use of null subjects diminished as lexical material in the AUX category emerged, a finding consistent with the observation Hyman (1986) had made for L1 acquisition. For this reason, Hilles concluded that modals and expletives both may work together as triggers for parameter setting and, in the case of L2 acquisition, resetting.

Hilles’ account has several obvious limitations. While longitudinal data may be very useful, the use of single subject makes generalizability of Hilles’ conclusions difficult. Furthermore, there is a great deal of subjectivity in taking English utterances, reconstructing how they ‘might’ have been said in Spanish, and then drawing conclusions based on that reconstruction. Finally, a raw count of missing subjects is not sensitive to discourse conditions in Spanish (or conversational English!) where subjects may have been properly deleted.

Phinney (1987) also began with Hyams’ (1983) hypothesis and attempted to empirically test it in a classroom environment through a contrastive study of L1 English
speakers learning Spanish and L1 speakers of Spanish learning English. Learners wrote compositions, which were then analyzed for the presence or absence of subject pronouns and the use of agreement morphology. Both groups performed well in regards to verbal agreement, but there were significant differences regarding the use of lexical and null subjects. L1 Spanish speakers omitted both expletive and referential pronouns in their L2 Spanish compositions, but L1 English speakers correctly omitted both types of pronouns. On the basis of this production data, Phinney (1987) concluded that the pro-drop parameter is reset easily and early for English learners of Spanish, but with greater difficulty for Spanish learners of English. Phinney then related her results to a theory of markedness, claiming that [+pro-drop] is the unmarked setting of the pro-drop parameter, making [-pro-drop] a marked setting. For Phinney (1987) this result meant that it was harder to change a parameter from an unmarked setting to a marked setting than it was to change from a marked setting to an unmarked setting.

Several questions remain unanswered in Phinney’s account. First, the two test groups composed their compositions under differing conditions—for one group the exercise was a test; for the other, the exercise was simply a class activity. It is uncertain what role the differing conditions may have played on the results. Second, as Phinney also notes, most of the forms were written in the first person, which may have skewed results. Third, Phinney’s results showed that the L1 Spanish speakers omitted subjects more frequently when there were errors in verb agreement, leaving open the question of the source of the errors. Finally, Phinney does not indicate whether subject pronouns were overt or covert in nontopic environments, or in those discourse situations where overt use of the pronoun is required also in Spanish.
Liceras (1988) provided additional production and grammaticality judgment tests of pro-drop and its properties. The 1988 study involved French and English learners of L2 Spanish. In this study, Liceras found no persuasive evidence that gaining a command of the Spanish inflectional system coincided with the various properties associated with pro-drop (e.g. inversion, optional or obligatory subject pronouns, or *that-trace* effects). She further found that even advanced learners ‘did not have native-like competence with respect to stylistic conventions that govern the use of inversion in Spanish’ (1989:115). Liceras (1988) concluded that different triggers may be required for different effects, with inversion possibly being triggered by the need in Spanish to have the preposition *a* before direct objects that refer to people, and *that-trace* possibly being triggered by the realization that empty complementizers are not permitted in Spanish relativization.

Liceras (1989) involved a grammaticality judgment of 17 items administered to 30 L1 English and 32 L1 French learners of Spanish from four levels of Spanish proficiency. The 17 items were manipulated to reveal information null subjects, overt expletives, inversion, and *that-trace*. The learners in this study were asked to make judgments, correct sentences with errors, and translate sentences into their L1.

From the results of these tasks, Liceras drew a number of conclusions: First, there was further confirmation of her 1988 finding that learners do not interpret *that-trace* sequences correctly, regardless of whether the complementizer is present or not, and that learners had more difficulties with inversion than with null subjects. Second, no lexical expletives in Spanish were accepted by learners. Most importantly, Liceras posited an implicational hierarchy regarding the order in which at least some of the properties associated with pro-drop are acquired. Liceras proposed that null subjects must be
acquired before inversion, and inversion before *that-trace*, at the same time granting that construction complexity and the structural properties of Spanish also play a role in the acquisition. Liceras noted that, if the Spanish setting for *that-trace* has been acquired, then inversion and null subjects have also been acquired, but the presence of inversion does not imply that acquisition of *that-trace* has taken place. The hierarchy of acquisition that emerges is: null subjects > inversion > *that-trace*.

This hierarchy would decompose the pro-drop parameter into different grammatical elements, the acquisitional order of which could then be empirically tested. If indeed the various grammatical phenomena that have been associated with pro-drop are acquired at different times, it would be reasonable to assume that there is no single parameter at all, but rather a progressive restructuring of a grammar in such a way that each new restructuring is evidenced by new grammatical effects. Liceras does not go so far as making this claim, but this insight is a key point of this dissertation.

In another study, Liceras (1989) combined grammaticality judgments with a translation task, providing a fuller insight into the language competence of the L2 learner; however, one complication of this study was that sentences often involved more than one feature, leaving some uncertainty as to what feature the reader was judging as grammatical or ungrammatical. This is a drawback to grammaticality judgment tests in which learners give only a binary response to the grammaticality of an item. In Liceras’ study, this problem is mitigated by including a correction and translation, but a clearer approach might be to provide subjects with clear choices varying only in the presence or absence of a targeted feature. An additional drawback to the Liceras (1989) study is the number of conditions that are spread out over a small number of test items. This resulted
in some properties of pro-drop being tested in only a couple of items. An improved methodology would propose more items per condition, and more total items.

Few SLA studies have provided true longitudinal data of the development of non-native Spanish speakers. Hilles (1986) attempted this but, as we saw, the study was limited to a single subject. Liceras (et al., 1997) represented another attempt to obtain longitudinal data from 5 high school and 11 university students acquiring L2 Spanish. Data collection took place through three recorded, 30-minute interviews eliciting spontaneous speech. The interviews occurred after students’ formal exposure to Spanish reached 50 hours, 65 hours, and 80 hours.

Liceras (et al.) evaluated the recorded data for the total number of sentences, the use of pro, the use of personal pronouns, and instances of incorrect morphology. The goal of this study was to evaluate several current hypotheses in grammatical theory as they relate to the status of null subjects and clitic pronouns: the minimal sentence hypothesis (Radford 1990), the short sentence hypothesis (Meisel and Müller 1992), the full sentence hypothesis (Hyams 1994, Rizzi 1994), and the VP-internal hypothesis (Zagona 1982, Koopman and Sportiche 1991). The details of each of their findings regarding each of these hypotheses are not significant here, and the results of their study were somewhat inconclusive, but Liceras (et al., 1997) is important for this dissertation in two regards: first, it provides another example of SLA research providing empirical tests of current grammatical theory; second, Liceras (et al., 1997:128) conclude that ‘...further research is necessary to define the various stages in the development of the non-native pronominal system’, and they suggest that a comprehensive account of non-native competence requires not only longitudinal data but also ‘...specific tests given at
different stages of the development of the non-native grammar’ (1997:128). This
dissertation seeks to make progress precisely in these areas.

Liceras and Díaz (1999) represents another attempt to account for the distribution
of null subjects and pronominal subjects in L2 Spanish using advances in grammatical
theory. Liceras and Díaz use proposals by Rizzi (1994) and Hyams (1994) to account for
pro, and they conclude that as non-native learners construct their grammars, they use
default licensing that allows null pronouns, provided the null pronouns can be identified.

Liceras and Díaz’ study looked at two different production data: (1) half-hour
interviews in which beginning learners, 12 year olds (n=5) and university students (n=6),
answered questions and were asked to create stories regarding characters and actions
depicted in a comic strip, and (2) narratives produced by advanced-intermediate, non-
native speakers (n=15) in which speakers were asked to tell a story based on one of their
favorite films. The first languages of the subjects in the second task were Chinese,
English, French, German, and Japanese. Three native speakers of Spanish were also
asked to perform both of the tasks. This study was ambitious in its breadth, with subjects
of differing ages, L1s, proficiency levels, and even tasks; however, given the relatively
small number of subjects, this amount of diversity may be more of a weakness than a
strength.

Liceras and Díaz found that even in early interlanguage learners produce null
subjects both in matrix and subordinate clauses. Using the theoretical insights of Rizzi
(1994) and Hyams (1994), they give an analysis where identification via discourse of f-
features overrides identification through subject pronouns. Liceras and Díaz believed this
to be a natural result for two reasons: (1) there are a large number of null subjects in the
Spanish input (saliency), and (2) learning the subjective, objective, and oblique inventories of Spanish pronouns is difficult (avoidance).

The type of production data provided here nicely complements a comprehension task, but it should be noted that Liceras and Díaz's study is representative of many other studies reviewed in this section in that it does not demonstrate how a parameter-setting framework accounts for developmental stages reflecting the interaction of discoursal and syntactic requirements. The binary nature of a parameter-setting framework does not easily account for stages in developmental route, and even Liceras and Díaz grant that interlanguages do not appear to be instantiations of coherent parametric options. Interlanguage evidence does not suggest that, for example, English second language learners of Spanish may simply ‘turn on’ a null subject parameter at some point. Rather, the evidence reveals just what a parameter-setting model should not: learners do not move directly from a 'non-pro-drop' stage to a 'pro-drop' stage; instead, there are clear developmental stages in which more than a single binary parameter is involved.

With the exception of a handful of studies (notably, Liceras 1988, 1989, Liceras and Díaz 1995, Pérez-Leroux and Glass 1997, 1999, Pérez-Leroux, et al. 1999), most L2 research on pro-drop has ignored semantic and discoursal interfaces with syntax in learners’ grammars. Many earlier studies operated with the premise that null-subjects are optional in languages such as Spanish and Italian, or simply described the conditions under which pro-drop is permissible. More recent L2 research (e.g. Pérez-Leroux et al. 1999, LaFond, Hayes, and Bhatt 2001) argues that such 'optionality' evaporates when concerns for information structure are included in the equation. Second language
learners face the challenge of learning that subjects are obligatorily absent in some contexts and obligatorily present in other contexts.

Pérez-Leroux, et al. (1999), demonstrate that interfaces between discourse and syntax in pro-drop languages control the distribution of null and overt pronouns. In Spanish, a null subject is permissible only when the antecedent is a topic; it is prohibited when the antecedent has nontopic status, as shown in 3.7 and 3.8 (Adapted from Pérez-Leroux, et al.1999): 9

(3.7.) a. ¿Quién canta?
   ‘Who sings?’

   b. Ellai/*Ø canta.
   ‘She sings.’

(3.8.) a. ¿Qué canta Cecilia?
   ‘What does Cecilia sing?’

   b. Ø/??Ella canta boleros.
   ‘She sings boleros.’

In 3.7, ella has not yet been activated in the discourse, prohibiting a null pronoun, but in 3.8, ella is noticeably worse than the null pronoun, because the discourse contains a recoverable 3sg feminine reference.

These examples also suggest that issues of focus may be involved in the distribution of null subjects. The answer to 3.7a must be focused, and when this is the case, null subjects are prohibited. Constraints related to focus play a significant role in the analysis used in this dissertation. As Grimshaw and Samek-Lodovici (1995) noted, subjects that undergo what was previously believed to be ‘free’inversion are not free at all; these subjects are actually focused in the discourse. Thus, concerns for information
structure relations may place conditions on inversion. If focus, like topic, interfaces with syntax in significant ways to determine the acceptability of sentences, then it, too can not be ignored in an account of L2 development.

Although this section’s review of the literature may seem to illustrate that there is little agreement on how the L2 acquisition of pro-drop proceeds, a certain set of facts regarding the acquisition of null subjects does emerge. First, despite the fact that null subjects appear early in L2 learners from a non-pro-drop language who are beginning to acquire a pro-drop language (Phinney 1987, Liceras and Díaz 1995), these learners initially overgenerate overt pronouns. This was noted early on by Fleming (1977), and confirmed by Liceras (1988), and Al-Kasey and Pérez-Leroux (1998). Second, despite their own overgeneration of overt pronouns in production tasks, L2 learners can sometimes detect the ungrammaticality of overgenerated overt pronouns, as shown in Liceras (1988). Third, there appears to be an implicational hierarchy regarding the order in which at least some of the properties associated with pro-drop are acquired (Liceras 1989). Finally, as Al-Kasey and Pérez-Leroux (1998) note, L2 learner errors are systematic, patterned, and related to the influence of the L1. If learners’ errors are systematic and patterned and there is an implicational hierarchy to the properties of pro-drop, then learners should follow a predictable developmental path as they acquire the pro-drop. The details of such a path have yet to be explicitly described. One of the contributions this dissertation makes to SLA research, in addition to an analysis offered within a particular theoretical framework (Chapter 6), is a clearer statement of the facts related to the path taken by L2 learners regarding pro-drop (Chapter 5).
3.6 SLA accounts of pro-drop from an OT perspective

This section completes the selected review of SLA literature by discussing two accounts related to pro-drop from an OT perspective: Park (2000) and LaFond, Hayes, and Bhatt (2001). Both of these studies begin to fill a gap between linguistic theorizing and its application to second language acquisition, but both of these studies have weaknesses that this dissertation attempts to address. Park's account is limited to null arguments in child second language acquisition, and as Park herself admits (2000:38) child L2 acquisition may have more in common with L1 acquisition than adult L2 acquisition. LaFond, Hayes, and Bhatt's account deals more directly with pro-drop from an adult SLA perspective, but this study is restricted to a consideration of null subjects, leaving further pieces of the pro-drop puzzle to later research. This dissertation supplies the further development of ideas prompted by this study.

For her study of null arguments (both subject and object) and wh-questions, Park (2000) used production data (collected by the National Center for Bilingual Research) from six Korean children learning English as a second language in a bilingual education school program. This data was part of a corpus of natural and elicited-interview speech. Park's goal was to examine whether two current theories, MP and OT, 'could account for language acquisition phenomena in real time' (2000:226), a goal that she concluded neither theory in its current state is prepared to handle.

According to Park, the feature-checking model of MP does not account for the distribution of null arguments. To make an MP account work, Park proposed that Korean has agreement features similar to Spanish-type languages, but that these features show up very late in children's grammars (past the age of six to eight years!), manifested by the
honorific marker *si* on the verb. Park concedes that such a claim involves quite a stretch for MP theory.

In regards to OT, Park used the constraints found in Speas (1997) and discussed in Chapter 2 of this dissertation. Once again, Park assumed a very late development of agreement features, perhaps as the result of inactivity of constraints related to agreement and phi-features. Park also argued that DROPTOPIC is a floating constraint, variably ranked along the constraint hierarchy. She notes that if this is the case, the placement of DROPTOPIC is more useful for explaining the distribution of null subjects in differing registers than in language development. Regarding this possibility, Park writes (2000:231):

This implies that OT may be more of a grammar of pragmatics than a grammar of syntax. Those who consider that pragmatics is not a part of linguistic competence may want to argue that OT is a grammar of linguistic performance.

Park reports that she considers pragmatics to be a module of linguistic competence, so she would not herself draw the conclusion that OT is a performance grammar.

As mentioned above, the value of Park’s account for adult SLA may be limited. Her bold proposal regarding agreement in Korean may spark further study for L1 studies of language development in Korean, but it admittedly falls short of providing a developmental account that is useful for adult L2 acquisition.

LaFond, Hayes, and Bhatt (2001) more directly addressed the issue of learning null subjects in a second language from an OT perspective. This study investigated the L2 acquisition of Spanish and Italian by L1 English learners at three universities: the University of South Carolina, the Pennsylvania State University, and the University of Arizona. A total of 210 subjects, divided into five proficiency levels for Spanish and four
proficiency levels for Italian, were given a written grammaticality judgment that investigated the use of null subjects in conversational dialogues.

Unlike most previous studies, these dialogues established a clear discourse context, and learners were asked to make a choice between two responses that best completed the dialogue. The responses represented syntactic minimal pairs, differing only in their presence or absence of a null subject. For some dialogues, the discourse context required a null subject; in others an overt subject was preferred. Error rates for topic and nontopic subjects were computed, and the results of this study revealed that very early learners initially undergenerate null subjects. At the next level of proficiency, learners use null subjects quite generally, without discriminating discourse context, resulting in overgeneration of null subjects. The appropriate subject retention in nontopic contexts was achieved only in advanced stages of acquisition. These results revealed a developmental path that LaFond, Hayes, and Bhatt interpreted from an OT perspective.

To do this, they begin by assuming the three constraints in 3.4 (from Grimshaw & Samek-Lodovici, 1995):

Table 3.2 Constraints implicated in the initial surfacing of null topics

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parse (PARSE)</td>
<td>Parse input constituents (Failed by unparsed elements in the input)</td>
</tr>
<tr>
<td>Subject (SUBJ)</td>
<td>The highest A-specifier in an extended projection must be filled (Failed by clauses without a subject in the canonical position)</td>
</tr>
<tr>
<td>Drop Topic (DROPT)</td>
<td>Leave arguments coreferent with the topic structurally unrealized (Failed by overt constituents which are coreferential with the topic)</td>
</tr>
</tbody>
</table>

Following the logic of OT, differences between English and Spanish regarding null subjects are the result in variance in the rankings of these constraints. The English
ranking, \textsc{Parse} \textgreater \textsc{Subject} \textgreater \textsc{DropTopic}, rejects candidates with null arguments Tableau 3.1:

Tableau 3.1 English grammatical system

<table>
<thead>
<tr>
<th>input: (&lt;\text{leave}(x), x = \text{topic}, x = \text{they}, y = \text{at eight}&gt;) tense = past</th>
<th>\textsc{Parse}</th>
<th>\textsc{Subj}</th>
<th>\textsc{DropT}</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (\emptyset) \textit{left at eight}</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. (\text{null})</td>
<td><em>!</em></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
| c. \(\text{They left at eight}\) | | | *
| d. \(\text{Left they at eight}\) | *! | * | *

In contrast, the Spanish ranking, \textsc{DropTopic} \textgreater \textsc{Parse} \textgreater \textsc{Subject}, selects candidates with null subjects as optimal when referencing a topic subject.

Tableau 3.2 Spanish grammatical system

<table>
<thead>
<tr>
<th>input: (&lt;\text{salir}(x), x = \text{topic}, x = \text{ellos}, y = \text{las ocho}&gt;) tense = past</th>
<th>\textsc{DropT}</th>
<th>\textsc{Parse}</th>
<th>\textsc{Subj}</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. (\emptyset) \textit{Salieron a las ocho}</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. (\text{null})</td>
<td><em>!</em></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. (\text{Ellos salieron a las ocho})</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| d. \(\text{Salieron ellos a las ocho}\) | *! | | *

LaFond, Hayes, and Bhatt discuss the developmental moves from the L1 English to the L2 Spanish or Italian. L1 English learners of Spanish initially transfer their L1 constraint ranking to the L2 data. Therefore, their grammars select a choice that is inappropriate for the L2. Following the Constraint Demotion Algorithm of Tesar and Smolensky (2000), the path of acquisition involves pairs of constraints being matched against each other, with the higher-ranked constraint (winner) being moved (demoted) lower than the loser-violated constraint. The initial demotion is shown in 3.3:
The resulting interlanguage hierarchy in Tableau 3.4 still does not converge on the L2 ranking; its choice is suboptimal:

Therefore, the learning algorithm applies once again and continues to apply as long as there is a mismatch between loser and winner marks (Tableau 3.5):

At the next stage, PARSE is demoted below DROPTOPIC:
Once this demotion occurs, the resulting hierarchy finally chooses the target optimal candidate, Tableau 3.7b:

Tableau 3.7  Interlanguage 2 (Target): DROP TOPIC » PARSE » SUBJECT

<table>
<thead>
<tr>
<th>input: subject=topic</th>
<th>DROP</th>
<th>PARSE</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP ellos [VP...]]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [IP ø [VP ...]]</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

The ranking in Tableau 3.7 converges on the ranking of Spanish, so the learning algorithm halts. The final restructured constraint ranking correctly reflects the observation that topics will be dropped; this ranking will necessarily retain nontopics.

This analysis given by LaFond, Hayes, and Bhatt accounts for the acquisition of null subjects, but does not yet explain how other properties associated with pro-drop languages fit into the picture. In the next chapter, the specific research questions and hypotheses that grow out of this development will suggest that inversion and that-trace effects may also be characterized in similar fashion as the null subject analysis of LaFond, Hayes, and Bhatt (2001). The null subject analysis of LaFond, Hayes, and Bhatt (2001) will also be given further refinement in light of the new data this dissertation provides. In so doing, this dissertation will provide a developmental account that provides a clear contribution to SLA research — it will offer an explanation of why certain pro-drop effects surface before others and why neither discoursal nor syntactic explanations can solely explain the path that learners take in their acquisition of Spanish.
Notes

1 The Audio-Lingual Method was the name given to the pronunciation and pattern drills used by the Army Specialized Training Program (ASTP) during World War II and based on the structural approach described by Fries (1948).

2 This ‘thought experiment’ provides evidence of another principle — that questions themselves indoctrinate and direct us to look at the world in specific ways.

3 In the context of second language acquisition, it might be interesting to pursue the flip side of Plato’s problem, what Chomsky has called ‘Orwell’s problem,’ i.e. how is it that human beings know so little given the amount of information to which they have access. Taken together, they lead us to ask why, given input that is so rich in some ways and so impoverished in others, humans acquire what they do.

4 ‘Atelic’ verbs are those that refer cumulatively, without a specific endpoint. In contrast, ‘telic’ verbs are delimited by a clear point at which the action of the verb is ends.

5 As in messages such as ‘Guess what? ∅ went to work today and Mary wasn’t there, so ∅ went downstairs and ∅ asked Sue if I could use her parking space...’

6 The lower number of French speakers was the result of White treating this group as a control group, since French (like English) is a non-pro-drop language. For those operating within a parameter-setting model, this assumption may be uncontroversial, but in a OT account such as that proposed in this dissertation, differences between French and English constraint hierarchies may challenge the grouping of French and English together in this manner.

7 X / (Y+X), where x=the number of instances of null subjects and y=the instances where null subjects would be predicted to have occurred in Spanish, but did not appear in English. Hilles did not include null expletives in this count because she thought possible triggers must be counted separately.

8 For example, Liceras et al. (1997:128) could not determine whether non-native productions of pronouns had the same value as native production of clitics because they could not determine ‘whether they are grammatical items or morphological parts of the verb’.

9 Liceras and Díaz (1995) hold that L2 learners’ null subjects are not necessarily instances of topic-drop. They argue that the learner’s production of null subjects is the result of local restructuring options (not reparameterization). As Pérez-Leroux and Glass (1997) note, if Liceras and Díaz are right, the overt/null alternation of subject pronouns may involve different aspects behaving independently.
Chapter 4
Research Methodology

4.0 Introduction

The discussion in the previous chapter made clear that, despite the extensive number of studies on pro-drop, numerous unanswered questions remain. Many previous studies were based on assumptions no longer current in linguistic theory, and none have provided a full developmental account of pro-drop from an OT perspective. Many accounts that have addressed developmental issues from a parameter-setting perspective have used a small number of subjects, considered only a single property of pro-drop, or asked for grammaticality judgments of sentences that lacked discourse contexts.1 This dissertation attempts to address some of the previous shortcomings through the use of broad-based tests constructed to reveal information about the interfaces between discoursal and syntactic components of grammar related to null subjects, inversion, and that-trace. The specific research questions and hypotheses these tests attempt to address and the research methodology used in these studies are the focus of this chapter.

This chapter is organized as follows: Section 4.1 introduces the research agenda that this study begins to address through the specific research questions and hypotheses pursued in this study. Section 4.2 discusses the motivation for using a grammaticality judgment task as the centerpiece for this study, the limitations and validity of a study based on grammaticality judgments, and a brief description of how the grammaticality judgment task in this study was constructed (using a translation task) and refined
(through a small pilot study). Section 4.3 discusses the research pool used for the translation task, the pilot, and the main study, and the precise procedures used for collecting the data. Finally, Section 4.4 discusses in greater detail the creation of the tasks and the predicted results.

4.1 Research questions and hypotheses

This dissertation initiates a research program focused on answering questions related to linguistic theory, interlanguage theory, and learning theory. In terms of linguistic theory, the goal of this dissertation is to shed greater light on how discourse and syntax conflict in English and Spanish, especially as relates to the properties typically associated with pro-drop, and to uncover implications these conflicts may have for the L2 learning of Spanish. In terms of interlanguage theory, the goal is to explain how types and rates of error for English learners of Spanish reveal tensions between syntactic and discoursal constraints. In terms of learning theory, the goal is to test a particular learning algorithm to see if it guides L1 English learners to converge on the L2 Spanish constraint hierarchies.

To begin to address these broad concerns, this dissertation asks the following specific research questions:

(Q1) Is the proposed implicational hierarchy of Liceras (1989) empirically verifiable?

(Q2) If Liceras’s hierarchy is verified, how should this hierarchy be understood in terms of OT constraint interactions? If the hierarchy is not verified, what modifications to it are required?
(Q3) What precisely are the relevant constraints that are involved in the mismatch between the way English and Spanish handle null subjects, inversion, and that-trace? What interactions between discourse and syntax are implicated by these constraints?

(Q4) Does the restructuring of these constraints as learners converge upon the target grammar find a natural explanation under the Constraint Demotion Algorithm of Tesar and Smolensky (2000)?

(Q5) Do learners eventually converge, in all respects, upon the target grammar? Are there pieces of the larger pro-drop phenomenon that remain resistant to acquisition?

In agreement with the Full Transfer/Full Access hypothesis of Schwartz and Sprouse (1996), this dissertation will assume that the initial state of the learner is a fully specified ranking of her native language. While this ranking is the one that is used in initial attempts to parse and produce utterances in the second language, the learner’s subsequent acquisition is guided by the UG operation proposed in Optimality Theory (full access). Schwartz and Sprouse’s hypothesis finds support in previous studies on pro-drop (e.g. Phinney 1987, White 1985, Liceras 1989) that claimed parameter resetting is difficult (due to L1 transfer), but possible (due to access to UG).

If parameter re-setting entails acquiring all the clustering of properties that pertain to a particular setting, then re-setting could be understood as automatization of all the properties that fall under that setting;² but if the acquisitional path of a pro-drop language does not involve the simultaneous appearance of the cluster of properties associated with pro-drop — that is, if a certain property is not obligatorily present at a particular interlanguage stage — then we have evidence for a process involving something other than the single setting of a parameter.
These research questions and set of assumptions have led to the formulation of four hypotheses:

(H1) As Liceras (1989) suggested, an implicational hierarchy will be found between the initial acceptability of null subjects, inversion, and *that-trace*.

(H2) ‘Initial acceptability’ will be distinct from ‘correct use’ of null subjects, inversion, and *that-trace*. The acceptance of inversion will not imply that the accurate use of null subjects has been acquired, nor will the acceptance of *that-trace* imply that the accurate use of inversion has been acquired.

(H3) Some L2 learners will eventually converge on native-like usage of null subjects, inversion and *that-trace*, but this convergence will come in only the late stages of L2 acquisition and as a result of a sensitivity to the discoursal constraints of the language.

(H4) The developmental path taken by L2 learners may be characterized as the interaction of discrete discoursal and syntactic constraints, and this interaction will confirm the operation of the Constraint Demotion Algorithm of Tesar and Smolensky (2000).

The first hypothesis is based on the assumption that the results of Liceras (1989) will be replicated: null subjects will surface before inversion, and inversion will appear before *that-trace*. This hypothesis would only be falsified if the acquisitional order demonstrates that *that-trace* surfaces before inversion or null subjects, or that null subjects appeared after inversion.

The second hypothesis implies that the implicational hierarchy of Liceras (1989) requires further refinement. It separates licensing from use, claiming that Liceras’s hierarchy applies only to the order in which these grammatical properties will find acceptance in the L2 grammar, but not necessarily the order in which native-like use of them will be acquired. This hypothesis would be falsified if the data shows null subjects
are used natively before inversion is acquired, or inversion is used natively before that-
trace is acquired.

The third hypothesis references the issue of ultimate attainment. Although L2 acquisition differs from L1 acquisition most dramatically in that L1 acquisition under normal conditions, unlike L2 acquisition, results in invariable success, the variable success of L2 learners should mean that at least some of the exceptional L2 learners in this study will achieve native-like competence regarding pro-drop. The first part of this hypothesis will be confirmed if some L2 learners do converge on the target grammar. If no learners in this study converge on the target L2, that result will falsify the hypothesis for the data set present in this study; however, this would not justify the conclusion that ultimate native-like competence is impossible, only that it was not achieved in this study. The second part of this hypothesis, that the convergence comes late in the acquisition process, follows previous findings (e.g. Galván 1998, LaFond, Hayes, and Bhatt 2001) that argued nuanced interactions between discourse and syntax result in the late acquisition of these properties. This second part of the hypothesis will be falsified if learners achieve early mastery of all pro-drop properties or if their mastery is unrelated to the acquisition of discoursal constraints.

The final hypothesis predicts that the developmental path taken by L2 learners may be characterized in OT terms as an interaction of discoursal and syntactic constraints and that this interaction will confirm of the Constraint Demotion Algorithm of Tesar and Smolensky (2000). This hypothesis fails (at least for the present analysis) either if no set of discoursal and syntactic constraints is found that can account for the results of the study, or if the restructuring of the constraint hierarchy that best accounts for the
developmental path is one that does not find a clear explanation using the learning algorithm proposed by Tesar and Smolensky.

### 4.2 The validity of grammaticality judgments

The centerpiece study of this dissertation is a cross-sectional grammaticality judgment task intended to provide insight into the state of learners’ competence at various stages of linguistic development. Intuitions concerning the grammaticality or ungrammaticality of a sentence form part of native speakers’ grammatical competence; in other words, native speakers have the ability to make solid judgments about both what can be said in a language and about what cannot be said. An important part of learning a second language is gaining the ability to determine the grammaticality of sentences in the target language. Therefore, grammaticality judgment tests have been widely used to investigate pro-drop (White 1985, Lakshmanan 1986, Phinney 1987, Liceras 1989, Yates 1990, Toribio, Roebuck, and Lantolf 1993, Pérez-Leroux and Glass 1997, 1999, Galván 1998, et al.).

The validity of grammaticality judgments has been challenged on several accounts. First, grammaticality judgments attempt to uncover grammatical ‘competence’, but, as Culicover (1997) mentions, the concept of competence itself involves a double idealization. Culicover writes the following (1997:11):

> We (falsely) treat all native speakers exactly the same and we assume that there is some stable and well-defined store of knowledge of language in the mind of the native speaker. The methodological assumption is that by ignoring differences between individuals and imprecision in the knowledge of individuals we may nevertheless discover something substantive and correct about natural language.
Not all researchers are willing to accept this idealization, and if the concept of ‘competence’ is questioned, any measure intended to arrive at that concept is also questionable.

Second, even among those who accept the common competence/performance distinction, it has been argued that grammaticality judgment tasks do not access the competence system in an untainted manner. The ‘performance’ of grammaticality judgments may itself vary, dependent on the mental state of the subjects making the judgment, the amount of time given for the task, or the ability of subjects to mentally construct possible scenarios permitting the utterance. For example, Schmerling (1978) argued that synonymy judgments (a type of grammaticality judgment) do not provide a useful test for an analysis of predicates taking infinitival complements. She argued that informants’ base their judgments on more than just their knowledge of the language, bringing into the process special strategies used to deal with example sentences, strategies which interact with the informants knowledge about the world as well as their language.

Schmerling looked at the contrast between 4.1 and 4.2, where the (a) and (b) examples in 4.1 are thought to be synonymous, but the (a) and (b) examples in 4.2 are not:

(4.1.) a. *I expected the doctor to examine John.*
    b. *I expected John to be examined by the doctor.*

(4.2.) a. *I persuaded the doctor to examine John.*
    b. *I persuaded John to be examined by the doctor.*

She notes that since Rosenbaum (1967), the distinction between these sentences has been accounted for by concluding that *persuade* sentences exhibit an Equi analysis while
expected sentences involve a Raising analysis, supporting the position that synonymy judgments reveal complex grammatical operations.

Schmerling shows that this type of analysis does not work well for a verb such as allow. Using a synonymy test, one could analyze allow as an Equi verb given the contrast in 4.3, but that analysis does not work if examples such as those in 4.4 are added, where allow looks like a Raising verb:

(4.3.) a. *I allowed the doctor to examine John.*
    b. *I allowed John to be examined by the doctor.*

(4.4.) a. *The new regulations allow there to be intolerable situations like this.*
    b. *The administration allowed unfair advantage to be taken of the strike.*

Schmerling argues that speakers have far more restricted direct intuitions of semantic meaning than previously thought — pragmatic factors have a way of slipping in to informants judgments of nonsynonymy. Informants decide, in absence of contrary pragmatic evidence, on the meaning of granting permission, as in 4.3, rather than the meaning do nothing to prevent, as in 4.4. This strategy is not a part of the grammar; it is, rather, a part of what it means to be an informant. Informants’ judgments here are based not only on syntactic knowledge but also on their own personal knowledge of the world.

This article calls into question the use of grammaticality judgments to gain a view into the underlying grammatical system, but there are several fundamental reasons to use them despite these objections. First, although underlying competence cannot be directly accessed, whatever access we do obtain through performance systems may potentially reveal information about the grammatical system of the learner. In other words, the inevitable fact of some variability in the performance of speakers even on grammaticality
judgment tests does not invalidate their use, though it should prompt us to be somewhat cautious about conclusions based on these judgments alone.

Second, the grammaticality task used in this study employs several measures designed to enhance its validity. For example, subjects in this study were not asked to make contextless judgments of grammaticality. Whereas traditional grammaticality judgment tasks frequently provide a single sentence and then ask subjects whether such a sentence is grammatical, the task in this study supplied a discourse context in the form of a dialogue. Additionally, subjects were not simply asked for a positive or negative judgment regarding a sentence; rather, they were given a choice between a pair of sentences that were minimally different in regards to the condition under investigation. This choice somewhat reduced the chance that subjects would reject a particular ungrammatical sentence for a reason unrelated to the grammatical property being studied.

Third, grammaticality judgments do not need to be used in isolation. Although this dissertation focuses on the results of a study that used grammaticality judgments, this was not the only measure used here. Production tasks used in concert with grammaticality judgment tasks provide additional windows into the learner competence (Phinney 1987, Galván 1998). In the current study, the results of a translation task confirmed and augmented the results of the grammaticality judgment task. The translation task involved translating from English into Spanish the items that would then be used to create the grammaticality judgment task. In addition, a pilot study of the completed grammaticality judgment task subjected the task to further scrutiny, to insure that subjects were not confused about the discourse settings presented in the dialogues.
Different subjects were used for the translation, pilot, and grammaticality judgment tasks. These subjects will be discussed in the next section.

Therefore, despite the drawbacks related to grammaticality judgments, their use (which, as mentioned above, has been the standard practice in SLA research of pro-drop) is justified here. The modifications that have been implemented and the additional production measure as a confirming test enhance the validity of the test used in this study.

### 4.3 Research subjects and data collection procedures

Research subjects for the translation task, the pilot study, and the grammaticality judgment task were drawn from a variety of venues that are described here. All research subjects were adult members of academic communities in South Carolina or Pennsylvania who volunteered to participate in this study. A total of 370 subjects participated in these studies: 124 completed the translation task, 39 completed the pilot, and 207 completed the grammaticality judgment task. Of the 370 subjects, 48 were native speakers of Spanish, 18 of whom were used for a control group for the translation task, and 30 of whom were used for the grammaticality judgment task. The native speakers of Spanish were from a variety of countries (Columbia, Costa Rica, Ecuador, Mexico, Peru, Puerto Rico, Spain, and Venezuela), a fact that was potentially important, since there are some differences in the Spanish spoken in these countries, but the results of the item analysis revealed no statistically significant differences related to country for the items in this study. The small pilot study did not use native speaker controls.

Subjects for the translation task were divided into four proficiency levels: beginning, intermediate, and advanced (based upon their placement in class levels and
number of years of Spanish instruction), and a fourth group consisting of native speaker controls. Subjects in the first three levels were native speakers of English who were learning Spanish at the University of South Carolina. The native speaker control group included native Spanish instructors, graduate teaching assistants in Spanish, and other graduate students at the University of South Carolina. There were 30 subjects in the beginning level, 36 in the intermediate level, 40 in the advanced level, and 18 native speakers as a control group. Participants in this study were given two dialogues and were asked to translate these into Spanish. Dialogues 6 and 8 (also found in Appendix A) serve as examples.

(6) Watching television

Julio: Iris, do you like watching sports on television.
Iris: Not too much, but I sometimes watch tennis.
Julio: Really? I like tennis too.
Iris: Have you been watching the US Open?
Julio: Yes, I especially like Todd Martin.
Iris: Who do you think will win the US Open?

(8) At the restaurant

Esther: This food is great.
Isabel: Yes. I like to come here.
Esther: Are you going to the meeting tomorrow?
Isabel: I’m going to decide this afternoon.
Esther: Who is going to be there?
Isabel: A group of women from Cuba will be there.

Subjects were given sufficient time to complete the dialogues; even at the beginning level most completed their translations in under 20 minutes. Although each participant translated two dialogues, they did not all translate the same dialogues. The specific
dialogues given to each participant were distributed randomly from a pool of 36 task items. All 36 tasks for this study may be found in Appendix A.

The 36 task items that were translated became the basis for the items used in the pilot and grammaticality judgment tasks. The pilot involved a 36-item grammaticality judgment task. The items were divided evenly between the grammatical properties of that null subjects, inversion and that-trace. The purpose for the pilot was to assess the time required to administer the test, and the clarity of the instructions, and the quality of the individual test items before they were used for the larger grammaticality study. This pilot was administered to 39 intermediate level learners of Spanish at the University of South Carolina. The results of the pilot were used to revise certain dialogues that were confusing to the subjects due to difficulties with the vocabulary, lack of clarity regarding the discourse context, or other unforeseen ambiguities in the text.

The revision of the pilot became the main grammaticality test that was administered to 207 subjects. Subjects in this larger study were divided into five proficiency levels: 64 beginners, 51 intermediates, 56 advanced, 6 near-native speakers of Spanish, and 30 native speakers of Spanish. Placement in these levels was done prior to any examination of the results, on the basis of Spanish class level and years of instruction. Of the L1 English learners of Spanish, 98 were students at the Pennsylvania State University and 79 were students at the University of South Carolina. The native speaker control group for this task was completely drawn from a pool of international students from Spanish speaking countries who were studying English as a second language in Columbia, South Carolina. These students were enrolled at the English Programs for Internationals (EPI), neighboring the University of South Carolina.
All subjects for the grammaticality judgment task were given the 36 dialogues in Spanish. An example dialogue is the Spanish translation of dialogue 8 (also found in Appendix B):

(8) *En el restaurante*

*Esther:* Esta comida está rica.
*Isabel:* Sí. A mí gusta venir aquí.
*Esther:* ¿Vas a la reunión mañana?
*Isabel:* Voy a decidir esta tarde.
*Esther:* ¿Quién va a estar allí?
*Isabel A:* Un grupo de mujeres de Cuba va a estar allí. □
*Isabel B:* Va a estar allí un grupo de mujeres de Cuba. □

For each of these dialogues, subjects were asked to choose the response that best completed the dialogue by placing a check after the ‘A’ or ‘B’ response. The complete test of 36 items is given in Appendix B.

The distribution of subject participants among levels are shown in Table 4.1:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation</td>
<td>30</td>
<td>36</td>
<td>40</td>
<td>-</td>
<td>18</td>
</tr>
<tr>
<td>Pilot</td>
<td>-</td>
<td>39</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gramm-Judgment</td>
<td>64</td>
<td>51</td>
<td>56</td>
<td>6</td>
<td>30</td>
</tr>
</tbody>
</table>

Several considerations went into grouping of proficiency levels in Table 4.1. First, since the grammaticality judgment task in this study is seeking to reveal, in part, whether native speakers of English can eventually converge on the L2 target language, the addition of a very advanced (i.e. near-native) group for this study appeared prudent. It is among near-natives, presumably, that the convergence on the L2 is most likely to occur, if it occurs at all. Second, since the translation task and the pilot were being used
in this study for the purposes of test creation and refinement, as well as to broadly confirm the patterns anticipated in the grammaticality judgment task, a near-native group was not as important for those tasks. This is, however, a limitation of the present study, and future research may wish to replicate both the translation task and the grammaticality judgment tasks, using more near-native speakers for both of the tasks.

Finally, some care is required in reporting the results of subjects from two different institutions. The proper placement of the subjects from the Pennsylvania State University, vis-à-vis the subjects from the University of South Carolina, was achieved through discussions with instructors from both institutions, the comparison of course syllabi and texts, and a comparison of demographic data given on a brief questionnaire (i.e. years of Spanish instruction, time spent in Spanish-speaking countries, etc.)

The procedure used for data collection was as follows: Subjects were given the test instruments with an attached consent form. This form provided them with the general purpose of the study (to examine the development of Spanish language proficiency among native speakers of English studying Spanish as a foreign language) and the procedures that would be used. Subjects were informed of potential risks (fatigue, boredom, frustration) and potential benefits (reading comprehension or translation practice, the opportunity to contribute to empirical research in second language development). They were also assured that their answers would be stored securely and confidentially, and reported anonymously. The complete consent form is found in Appendix C.

Subjects were asked to sign and date the consent form, as well as to answer a brief questionnaire gathering demographic data about their native language, the amount of
Spanish instruction they had received, the amount of time they may have spent in
Spanish-speaking countries, etc. The information was used to cull from the total sample,
prior to an analysis of the results, those tests whose results may have been skewed due to
extraneous variables. For example, some subjects reported native languages other than
English or an inordinate number of years of instruction for the level of class they were
taking. Some also failed to properly complete the task properly (e.g. by checking more
than one box or skipping a page) so these tests were also culled from the total tabulations
used for analysis.

4.4 Task items and predicted results

This section discusses in greater detail factors that were taken into consideration
in the creation of the tasks, how this data was coded, and the predicted results. The same
goal was being pursued in each of this dissertation’s tasks—to reveal information about
the status of null subjects, inversion, and that-trace at various stages in the development
of learner grammars. Therefore, this section does not focus on the three different tests,
but rather on the three grammatical properties and the items created to investigate them in
all three tests.

A number of general considerations went into the creation of all the items used
here. Since the same test would be administered to all levels, the vocabulary needed to
carefully controlled to avoid lexical items that would be unfamiliar to beginning learners.
The vocabulary used was drawn from beginning texts in consultation with instructors of
beginning Spanish. Also, complex or advanced verbal structures (e.g. the subjunctive or
perfect tenses) were avoided, and sentence structure was kept as simple as the grammatical properties under investigation allowed.

The 36 items in all three tests were designed to investigate judgments regarding three conditions: null subjects, inversion, and *that-trace*. Each condition was further subdivided into subconditions. For example, items testing null subjects were of two types: those that tested the dropping of topic subjects and those that tested the retention of nontopic subjects. There were also two types of inversion items: those that involved presentational or contrastive focus and those that lacked this special discoursal feature. *That-trace* items in this study mostly contrasted sentences with a complementizer *que* with sentences lacking the complementizer, although two additional items tested whether judgments regarding [*+que/-null subject*] would contrast with [*+que/+null subject*]. Given the diverse set of grammatical properties and conditions, no distractors were included in the test design.

The test items created to investigate null subjects followed the methodology of LaFond, Hayes, and Bhatt (2001), with minor modifications. Two types of items were constructed to investigate null subjects. The first type is exemplified in 4.5a (dialogue 20 in Appendix B), translated in 4.5a':

(4.5.) Dialogue selecting a topic subject

a. *En la universidad*

*Simón:* Fue muy difícil esa clase de biología.
*Adriana:* Ah, ¿sí? ¿Qué estudias?
*Simón:* Historia. No me gustan las ciencias.
*Adriana:* Qué lástima. Las ciencias pueden ser muy interesantes.
*Simón:* ¿Cómo te interesaste por las ciencias?
*Adriana A:* Tuve una maestra buenísima en la escuela secundaria. √
*Adriana B:* Yo tuve una maestra buenísima en la escuela secundaria.
a'. At the university

Simón: That biology class was very difficult.
Adriana: Really? What is your major?
Simón: History. I don’t like science.
Adriana: That’s too bad. Science can be interesting.
Simón: How did you get interested in science?
Adriana: I had a very good teacher in high school.

The dialogue in 4.5 exemplifies a discourse setting that selects a phonetically unrealized topic subject. In 4.5, Adriana has been asked about her interest in science. Since Adriana and her interests are topics under discussion, and since Adriana is contextually recoverable as a discourse participant, a null subject is required in the response to Simón’s question. The final choice that omits the overt first person pronoun (yo) is, therefore, the preferred choice. Learners of Spanish who are sensitive to the null subject requirements of Spanish should identify the need for a null subject in these types of examples, but early learners (whose L1 English grammatical constraint rankings demand that sentences have overt subjects) are predicted to less reliably choose the null subject in these contexts.

In contrast, certain discourse contexts in Spanish require overt subjects, particularly when the subject is not easily recoverable or is not the topic of the discourse. An example is provided in 4.6a, (dialogue 17 in Appendix B), translated in 4.6a’:

(4.6.) Dialogue selecting overtly-realized subject

a. Después de una visita al museo

Luis: Fui al museo esta mañana.
Rosa: ¿Y qué viste?
Luis: Muchas pinturas de Picasso.
Rosa: A mí me gusta mucho Picasso.
Luis: A mí también; y por eso compré un póster en la tienda de regalos.
Rosa A: Yo también compré uno la semana pasada. √
Rosa B: También compré uno la semana pasada.
a’. After a trip to the museum

Luis: I went to the museum this morning.
Rosa: What did you see?
Luis: Lots of paintings by Picasso.
Rosa: I like Picasso a lot.
Luis: Me too; so I bought a poster in the gift store.
Rosa: I also bought one last week.

In 4.6, even though Rosa is a participant in the conversation and, thus, a contextually recoverable referent, she is not the topic of the discourse or the preceding utterance. Therefore, the retention of the pronoun yo is warranted. Although Spanish and English do not differ regarding the requirement for subjects in such sentences (i.e. both grammatical systems require overt subjects), LaFond, Hayes, and Bhatt (2001) have shown that early L2 Spanish learners sometimes overgenerate null subjects in such sentences, indicating that the interlanguage grammar of these learners is different than both the grammar they know (their L1) and the target grammar they are attempting to learn (the L2). Under the assumptions of the theoretical framework used in this study, it could be predicted that the output of learner grammars may differ from both the L1 and L2 until constraint rankings converge on the target language ranking that properly identifies in which contexts null subjects should appear and in which they should not. Therefore, even though both Spanish and English require the use of overt subjects in sentences such as 4.6, some variance between native and learner groups for such items is expected.

Test items created to examine inversion were fashioned to investigate whether the conclusions of Grimshaw and Samek-Lodovici (1995) and Samek-Lodovici (1996) for Italian could also be extended to Spanish. Those previous studies observed that, while
English does not permit the appearance of post-verbal subjects, Italian does. More importantly, inversion in Italian is not as free as is sometimes assumed — subjects obligatorily appear at the right edge of the VP when they are contrastively focused. The inversion items in the current study investigated whether discoursal requirements govern focus also in Spanish. Since, in English, syntactic constraints prohibit post-verbal subjects, English uses other means to indicate focus (e.g. by marking the prominent element with intonational stress); therefore, transfer effects of the L1 may be predicted in early learners, resulting in lower rates of acceptance of inverted subjects.

An example is provided in 4.7a (dialogue 4 in Appendix B), translated in 4.7a', of a context that would suggest contrastive focus of a subject:

(4.7.) Discourse context selects inverted subject

a. Asumiendo responsabilidad

Pablo: ¿De qué te ríes?
Janet: Mamá y Papá van a llegar pronto.
Pablo: ¡Ay! ¡La casa está muy sucia!
Janet: Te dije: ‘No invites a tus amigos.’
Pablo: ¿Me ayudas a limpiar la casa?
Janet A: ¡No! Tienes que limpiarla tú, no yo.
Janet B: ¡No! Tienes que limpiarla, no yo.

a'. Taking responsibility

Pablo: What are you laughing about?
Janet: Mom and dad are coming home early.
Pablo: Oh no! The house is very dirty!
Janet: I told you, ‘Don’t invite your friends.’
Pablo: Will you help me clean the house?
Janet: No! You will have to clean it, not me.
The dialogue in 4.7 presents a classic example of contrastive focus. In the dialogue, a brother and sister argue over who should clean the house. If contrastive focus calls for inversion in Spanish, Janet’s ‘A’ response should be preferred. This is in contrast to items where the subject is not as clearly focused. Consider example 4.8a (dialogue 22 in Appendix B), translated in 4.8a’:

(4.8.) Discourse context does not require inversion

a. Hablando por teléfono celular

- Armando: ¿Bueno?
- Patricia: Se fue un poquito antes que yo.
- Armando: Te veo pronto.
- Patricia A: Acabo de salir de la escuela yo.
- Patricia B: Yo acabo de salir de la escuela. √

a’. On a cell phone

- Armando: Hello?
- Patricia: Hi, Armando. I’m on my way home.
- Armando: Erica called. She is on her way too.
- Patricia: She left a little before me.
- Armando: See you soon.
- Patricia: I have just left the school.

Example 4.8 involves a cell phone conversation where Patricia calls Armando to let him know she is on her way home from school. When Armando says ‘See you soon’, Patricia replies that she has just left the school. The first person pronoun, yo, implies no special focus, contrastive or otherwise, and would not be expected to be inverted. It should be mentioned that in certain dialogues like these, native speakers would likely use a null subject (since the subject is a topic), so that neither the ‘A’ nor the ‘B’ response would be the most favored choice. Nevertheless, given the two lesser choices of either the overt
realization of the pronoun, or the overt and inverted realization of the pronoun, speakers should prefer the non-inverted choice, since the inverted choice should involve a greater degree of deviation from the expected norm.

The test items created to investigate *that*-trace effects manipulated the appearance of the complementizer *que* ‘that’. English permits the exclusion of the complementizer *that* in sentences where Spanish would require it (4.9a), and it prohibits the use of the complementizer *that* in sentences where Spanish would require it (4.9b). Furthermore, when English uses the complementizer, in contrast to Spanish, it requires the use of the subject pronoun following it (4.9c)

(4.9.)  
a. *Does he think (that) he will win the U.S. Open?*  
b. *Who, do you think that he will win the U.S. Open?*  
c. *Do you think that he/*∅ will win the U.S. Open?*  

The test items in this study explore these effects by providing dialogues in which sentences with a complementizer *que* are contrasted with sentences lacking the complementizer, exemplified in 4.10a (dialogue 6 in Appendix B), translated in 4.10a’, and by a few additional sentences that tested other manipulations of *que* and subjects. For example, two sentences contrasted [+*que*/+null subject] with [+*que*/-null subject], as in example 4.11a (dialogue 18 in Appendix B), translated in 4.11a’), where a null topic subject is required following the complementizer *que*.  

143
(4.10.) Que vs. no que

a. Mirando la televisión

Julio: Iris, ¿a ti te gusta mirar deportes en la tele?
Iris: No me gusta mucho, pero a veces veo el tenis.
Julio: ¿De veras? A mí me gusta el tenis también.
Iris: ¿Has estado mirando el U.S. Open?
Julio: Sí, me gusta especialmente Todd Martin.
Iris A: ¿Quién piensas va a ganar el U.S. Open?
Iris B: ¿Quién piensas que va a ganar el U.S. Open? ✓

a'. Watching television

Julio: Iris, do you like watching sports on television?
Iris: Not too much, but I sometimes watch tennis.
Julio: Really? I like tennis too.
Iris: Have you been watching the US Open?
Julio: Yes, I especially like Todd Martin.
Iris: Who do you think will win the US Open?

(4.11.) Que + null subject vs. que + overt subject

a. Carmen y Felipe en la oficina

Carmen: ¿Alguien me llamó cuando yo no estaba aquí?
Esther: Sí. Luis Pérez y Lilia Enríquez.
Carmen: Ok. ¿Alguien más?
Esther: No, no llamó nadie más.
Carmen: ¿Dijeron Luis y Lilia lo que querían?
Esther A: No, pero dijeron que iban a volver a llamar.
Esther B: No, pero dijeron que iban a volver a llamar. ✓

a’. Carmen and Felipe at the office.

Carmen: Did anyone call me while I was out?
Esther: Yes. Luis Pérez and Lilia Enríquez.
Carmen: Ok. Anyone else?
Esther: No, no one else called.
Carmen: Did Luis and Lilia say what they wanted?
Esther: No, but they said they are going to call again.

Of the three properties under investigation in this study, that-trace appears to be the most purely syntactic. Whereas the presence or absence of null subjects and
inversion are predicted to correspond to certain discourse conditions in Spanish, sentences with *that-trace* effects are not expected to display this same type of discourse sensitivity; however, as with null subjects and inversion, there may be some lag for learners between acceptance of the property and the correct use of the property. For example, learners are predicted to accept sentences displaying *that-trace* sooner than they reject ungrammatical sentences on the basis of the lack of the complementizer.

In summary, for all three grammatical properties, it is predicted that the test items in this study will reveal significant differences between native speakers and early learners, and that these differences will be the result of conflicts between the L1 and L2 grammatical systems of the learners. As learners progress toward the resolution of these conflicts, understood in this study as the reranking of constraints, they are expected to converge on the target grammar. If this prediction is correct, at the higher levels of study, few differences will be found between the way native speakers and advanced learners treat null subjects, inversion, and *that-trace*. The experimental results of the translation and grammaticality judgment tasks in this study are presented in the next chapter.
Notes

1 One study (Galván 1998) that was not a part of the selective review in the last chapter did use a larger number of research subjects: Galván administered a grammaticality judgment task and a written composition task to 222 native English-speaking learners of Spanish to uncover the use of null and lexical subjects in various syntactic and semantic-pragmatic environments. Galván’s parameter-based analysis concludes that learners do not begin their acquisition of Spanish with the L1 settings of the null subject parameter, and he disagrees with the claims of Liceras (1989) and Phinney (1987) that the default setting of the parameter is [+pro-drop]. Most importantly, Galván argues that the distribution of null subjects in Spanish is constrained by semantic, pragmatic, and syntactic factors.

2 If the use of null subjects is acquired before inversion and that-trace, then it is possible to claim that pro-drop is unrelated to inversion and that-trace, raising the question of why they should be studied together; however, this cluster of properties frequently (though not always) arise in certain languages, creating an observable correlation between properties x, y, and z. The reason for this correlation is a matter of interest. This dissertation has no disagreement with the use of the term ‘pro-drop’ as a shorthand to refer to this correlation, but the argument here is that pro-drop is not best conceived of as a parameter, because this clustering of properties is neither required (language may have one, or two, of the properties but not the third) nor simultaneous.

3 The issue of ultimate attainment has been vigorously debated. For a summary of the issues involved on each side, the reader is referred to the study by Coppieters (1987) and the response to it by Birdsong (1992).

4 ‘Equi’ refers to a syntactic operation hypothesized by early transformational grammar. This operation deleted a noun phrase when another identical noun phrase was present in the sentence; for example, I asked John to come, arose from the deletion of the second John in a pair of sentences, I asked John [John come]. This was thought to differ from ‘Raising’, a process of moving a noun phrase from the subordinate clause into the structure of the higher clause, as in I believe him, [t, to be honest].

5 In all cases, different subjects were used for the translation and grammaticality judgment tasks. Some of the subjects who participated in the translation task also participated in the pilot study. Since the pilot was being used solely to refine the reliability and validity of the grammaticality judgment task (i.e. no specific results are being reported from the pilot), the overlap between these two tasks poses no problem.
6 Intermediates were chosen because it was thought that their responses (being the middle group of the study) would permit the best average assessment of these items. In retrospect, refinement of the final task would have been better served by adding beginners and native speakers to the pilot; nevertheless, as the next chapter will discuss, the pilot did provide useful information for the creation of the grammaticality judgment task.

7 LaFond, Hayes, and Bhatt (2001) used dialogues similar to the ones used in this study, but the dialogues in that study did not provide titles for the dialogues or names for the participants in the dialogue. Dialogue participants in that study were simply labeled participant ‘A’ and participant ‘B’. The addition of titles and names in this dissertation’s study provides additional contextual information for each item, helping readers to access the ‘narrative schema’ (Chafe 1994) or ‘domain of action’ (Clark 1996) that would be appropriate for interpreting the conversation, thus giving a more natural basis for making discourse judgments.
Chapter 5

Results and Interpretation

5.0 Introduction

This chapter presents the results obtained from the translation task, the pilot task, and the main grammaticality judgment task. Section 5.1 discusses the coding and data analysis procedures used for the translation task, and it presents the results of this task. Analyzing the translation task posed several challenges, which are discussed in this section. The statistically valid findings of this test are very general, requiring the further refinement that the grammaticality judgment task yielded. Section 5.2 reports on the result of the pilot, focusing mostly on the qualitative issues specific to the purposes of the pilot. Section 5.3 discusses the coding and data analysis procedures used for the grammaticality judgment task and then presents the statistical results of this task. The significant differences between the responses of the various proficiency levels are discussed, organized around the three grammatical properties of null subjects (Section 5.3.1), inversion (Section 5.3.2) and that-trace (Section 5.3.3). Section 5.4 interprets these results in light of the implicational hierarchy of Liceras (1989), the issue of ultimate attainment, and the impact of discourse on grammatical choices, leading into the analysis provided in the next chapter. Finally, Section 5.5 briefly discusses limitations of the study revealed by the results.
5.1 Translation task

In the last chapter, we noted that the translation task involved participants from four proficiency levels (beginning, intermediate, advanced, and native speakers) and that each participant was asked to translate two dialogues. Ideally, for the purpose of statistical analyses, it would have been better to have each item translated by each subject, but the time involved in having a single subject translate 36 dialogues would be unreasonable; therefore, each subject (n=124) translated only two dialogues, yielding a total of 248 tokens for analysis. Each of the 36 items was translated at least twice by each proficiency level. These translation tokens were then coded with respect to the preferred responses by native speakers. For example, if native speakers consistently used a null subject in the translation of a particular sentence, then translations by other proficiency groups that lacked a null subject in that context were marked with a ‘1’, indicating a dispreferred translation. If the subject translated the sentence using the preferred translation, it was coded with a ‘0’ to indicate no variance from the native speaker norm; therefore, the coding yielded binary responses (preferred/dispreferred).

To model the binary responses here, the first attempt was to code them as ‘logits’, or ‘logs of odds ratios’. An odds ratio measures the probability of a preferred response versus the probability of a dispreferred response under two different sets of conditions (e.g. that-trace and inversion). The analysis was intended to be a logistic regression with two factors. The responses were not independent, since there are two responses from each subject, so it was necessary to include ‘subject’ in the model as a clustering variable.
(or random effect) to induce correlation among the responses for each subject. To do this, the SAS statistical package was used, employing PROC GENMOD.

Analysis of the items for this test proved problematic. This repeated measures model treated the responses of each subject as correlated data, not independent data. In all of the models used for these studies, in which subjects responded to multiple items, it was assumed that the items are correlated within the subject. The initial attempt was to analyze item and group as factors. For the binary response model, however, if any of the items or any of the groups had only one of the two responses present for all subjects who tested on that item or belonged to that group (for example, if for item 16, there were no ‘1’s, hence only ‘0’s), then the algorithm failed to converge. There were many such items in this case, and no further attempt was made to include item as a factor.1

Although the translation task did not reveal significant information regarding items, it did reveal that groups differed significantly in their response patterns. Table 5.1 shows the analysis of initial parameter estimates related to each group.

Table 5.1: Analysis of initial parameter estimates for translation task, by group

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Wald 95%</th>
<th>Confidence Limits</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning</td>
<td>1</td>
<td>-3.5041</td>
<td>0.6376</td>
<td>-4.7535</td>
<td>-2.2544</td>
<td>30.20</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Intermed.</td>
<td>1</td>
<td>-2.7675</td>
<td>0.6377</td>
<td>-4.0174</td>
<td>-1.5175</td>
<td>18.83</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Advanced</td>
<td>1</td>
<td>-1.4077</td>
<td>0.6627</td>
<td>-2.7065</td>
<td>-0.1088</td>
<td>4.51</td>
<td>0.0337</td>
</tr>
<tr>
<td>Native</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.1 sets the native speaker group as the baseline. The ‘DF’ designator in column 2 refers to the degrees of freedom, or the number of independent pieces of information that go into the estimate of a parameter. The test permitted learner groups to vary from the native group by at most one degree, because for each item the evaluation
was binary, based on whether the translation for an item did or did not use a null subject, inversion, or *that-trace*. The remaining columns show that native speaker response patterns are distinct from all other groups, including the advanced group (although evidence for the difference between advanced and native is marginal). The parameters are listed as log odds ratios. The value of $-3.5041$ in column 3 for the beginning group indicates that the odds of a preferred response for the native group is $\exp(3.5041) = 33$ times greater than the odds of a preferred response for the beginning group. In other words, natives use the grammatical constructions under investigation much more frequently than beginners. The bounds on these odds are obtained from the Wald 95% Confidence Limits ($\exp(2.2544), \exp(4.7538) = (9.5, 166)$). We may be 95% confident that the odds of a native speaker giving a preferred response are 9.5 to 166 times greater than that of a beginning speaker. Intermediate speakers were 15 times less likely to give a preferred response than native speakers. The Wald 95% Confidence Limits for this group ranged from 4.5 to 55.5 times less likely to give a preferred response.

The difference between native and advanced groups did not achieve significance under this test, largely because the Wald limits ranged from 1.1 to 14.9 times less likely for the advanced speakers to give a preferred response. The performance of the advanced group was, however, distinct from the native group, with $p = 0.0337$, and an odds ratio 4 times less likely than natives to provide the preferred response.

The results of the translation task reveal that the proficiency groups tested here do vary in regards to their ability to translate dialogues in a native-like manner with respect to null subjects, inversion, and *that-trace*. This study also revealed progressive movement towards native patterns, with the advanced level as a group differing from
native speakers in a less than significant way. Nevertheless, this study was not designed to indicate the exact loci of the differences in regards to items.

5.2 Pilot study

As discussed in the previous chapter, the purpose of the pilot was to assess the time required to administer the test, check the clarity of the instructions, and uncover difficulties related to vocabulary choice or ambiguities in the discourse context — in other words, to refine the test that would be used for the main study. Since numerous items were changed as a result of the pilot, and a comparison of different versions of the same item would not be valid, a separate statistical package was not run on the pilot. A description of the results, however, is reported here to give a sense of the directionality of responses, along with qualitative results related to the purpose of the task.

Subjects for the pilot study asked very few questions after the form was read and they appeared to clearly understand the directions. Less than 2% of the forms were inaccurately completed (i.e. most subjects correctly checked one and only one box, and did not skip any items). In a few instances, subjects began the test but chose not to finish it. These tests were excluded from the sample. There were also several tests turned in by subjects who reported a language other than English as their L1. These tests were also excluded from the sample. On a number of the tests, subjects marked vocabulary items that they did not understand. These items were replaced with more common vocabulary in the revised test. Following the pilot, two instances of advanced verbal forms were found, as well as a few typographical errors. All of these were corrected in the revised version.
The intermediate learners in the pilot normally selected null subjects when given the choice between a null and overt subject. This was true even for dialogues where an overt subject is preferred by native speakers.\textsuperscript{2} For example, for dialogues 17 and 28 below, the preferred native response is the ‘A’ choice, identified by a check mark, and native speakers made this choice about 90% of the time. In contrast, intermediate level subjects chose the ‘A’ response only about 50% of the time.\textsuperscript{3}

(17) \textit{Después de un paseo por el museo}\textsuperscript{4}
Luis: Fui al museo esta mañana.
Rosa: ¿Y qué viste?
Luis: Muchas pinturas de Picasso.
Rosa: A mí me gusta mucho Picasso.
Luis: A mí también; y por eso compré un póster en la tienda de regalos.
Rosa A: Yo también compré uno la semana pasada. √
Rosa B: También compré uno la semana pasada.

(28) \textit{Haciendo una acusación}
Carmen: ¿Entonces, qué está diciendo?
Rosa: Juana y usted estaban ahí en ese momento.
Carmen: ¿y?
Rosa: Que Juana y usted tenían una razón para hacerlo.
Carmen: Otro lo hizo.
Rosa B: Lo hicieron. Nadie más.

Intermediate speakers also registered a strong preference for non-inverted choices, and, again, registered a roughly 50% split even on items for which the preferred native choice was to invert nearly 97% of the time, as in dialogue 4 below:

(4) \textit{Asumiendo responsabilidad}
Pablo: ¿De qué te ríes?
Janet: Mamá y Papá van a llegar pronto.
Pablo: ¡Ay! ¡La casa está muy sucia!
Janet: Te dije: ‘No invites a tus amigos.’
Pablo: ¿Me ayudas a limpiar la casa?
Janet A: ¡No! Tienes que limpiarla tú, no yo. √
Janet B: ¡No! Tienes que limpiarla, no yo.
The pilot also revealed that intermediate learners had great difficulties with the correct distribution of *que* ‘that’ in these dialogues. For example, in dialogues such as 1 and 19 below, native speakers had a strong preference for the ‘B’ response (87% and 84%, respectively), but the intermediates in this pilot chose ‘B’ only 13% of the time for item 19 and not at all for dialogue 1.

(1) *Después de la matrícula en la escuela*

Alicia: ¿Te has inscrito en tus clases?
Marta: Sí, pero tengo un problema.
Alicia: ¿Qué pasó?
Marta: No pueden encontrar mi cheque.
Alicia: Se han equivocado con mi cuenta también.
Marta A: ¿Quién piensas nos puede ayudar?
Marta B: ¿Quién piensas que nos puede ayudar? √

(19) *En el aeropuerto*

Rosalía: ¡Hola! Roberto.
Roberto: Bienvenida.
Rosalía: Estoy muy cansada.
Roberto: No tendrás mucho tiempo para descansar.
Rosalía: ¿Por qué? Voy a casa.
Roberto A: ¿Quién esperas va a trabajar esta noche?
Roberto B: ¿Quién esperas que va a trabajar esta noche? √

Intermediates were willing to accept *that-trace* violations under certain conditions. For example, for dialogues 14 and 36 (below) intermediates chose the *that-trace* violation at rates of 61% and 69%, respectively.

(14) *Una mujer habla por teléfono con su marido*

Carmen: ¿Hola?
Felipe: Buenos días, Carmen.
Carmen: ¿Visitaste al médico? ¿Tienes noticias?
Felipe: Sí, esta mañana.
Carmen: ¿Qué te dijo?
Felipe A: Que no sabe nada todavía. √
Felipe B: Que él no sabe nada todavía.
Interestingly, the pattern of responses here suggests that intermediate learners’ difficulty with the distribution of *que* in Spanish may not rest with the violation of *that-trace* but with identifying sentences where *que* is required. This is most clearly seen in dialogues such as 14. This also seems to hold for dialogue 36; however, it should be noted that in that dialogue learners may have been reacting more against the inverted subject than for the *that-trace* sequence.

To summarize the results of the pilot: subjects in this study appeared to have acquired null subjects, but lacked a sensitivity to discourse conditions that may regulate their use. They displayed dispreference for inversion, preferring non-inverted orders even for sentences where native speakers prefer inversion, although they did select inversion more often when the inverted item was focused. Intermediates had problems with sentences involving *that-trace* violations, but it is not clear from this study whether this problem is one of possessing a grammar prohibiting *that-trace* or whether the difficulty is rather one of not recognizing the obligatoriness of *que* in particular contexts.

### 5.3 Grammaticality judgment task

This section reports on the grammaticality judgment task which, as noted previously, consisted of learners making binary choices on 36 items differing in the use
of null subjects, inversion, and *that-trace*. There were two forms of this test used for the main study; each presented items in different random orders. Analysis of the results of the two forms of the test found no statistical difference between them, so the order of the presentation of the items was not significant.

The administration of the pilot revealed that the test would take 15–20 minutes for intermediate learners to complete. It was then assumed that, for the main grammaticality judgment task, beginning learners would take slightly longer, and advanced learners slightly less time. This proved to be correct. Advanced speakers and above usually completed the test in 10–15 minutes, while beginners normally took 20–30 minutes. A common consent form was used for all levels, so as a result of the pilot, subjects were all told that the test should be completed in 20–30 minutes. The consent form was refined in two additional ways. First, the original did not inform subjects that as they completed the test they might experience some fatigue, boredom, or frustration. These reactions were observed in the pilot group, but the presence of the advisory in the subsequent tests appeared to mitigate these problems.

A total of 207 subjects responded to each of the 36 dialogues, yielding a total of 7452 tokens for analysis. Each of these tokens was coded with a ‘0’ if the respondent chose choice ‘A’ and a ‘1’ if the respondent chose choice ‘B’. These responses were then analyzed compared to the baseline responses of the native group. An analysis of initial parameter estimates (Table 5.2) yielded results similar to those of the analysis of the translation task:
Table 5.2: Analysis of initial parameter estimates for gramm. judgment task, by group

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Wald 95%</th>
<th>Confidence Limits</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning</td>
<td>1</td>
<td>-0.5142</td>
<td>0.0784</td>
<td>-0.6678</td>
<td>-0.3606 43.06</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Intermed.</td>
<td>1</td>
<td>-0.3428</td>
<td>0.0814</td>
<td>-0.5024</td>
<td>-0.1832 17.72</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Advanced</td>
<td>1</td>
<td>-0.3649</td>
<td>0.0801</td>
<td>-0.5219</td>
<td>-0.2079 20.75</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Near-Nat.</td>
<td>1</td>
<td>-0.5030</td>
<td>0.1551</td>
<td>-0.8070</td>
<td>-0.1990 10.52</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>Native</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000 - -</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2 shows that native speaker response patterns were again significantly distinct ($p < .0001$) from beginning and intermediate groups; this test additionally showed significance between the advanced and native groups ($p < .0001$) and a clear, but not significant, difference between the near-native group and the native speakers ($p = 0.0012$).

The log odds ratios in the grammaticality judgment task were smaller than in the translation task, but the larger number of respondents and the fact that each subject responded to each item still permitted valid analysis of significances.

The grammaticality judgment task shared one problem with the translation task in that running a global model to analyze item and group parameters simultaneously did not work due to some groups always getting the same number on an item. Fortunately, there are other default methods that can be used to test for item–by–group differences. The standard way of checking whether there is a group effect on a particular item is to use a Fisher Exact Probability Test, a useful alternative to chi-square when that test would not yield valid results. The Fisher test involves a more complicated formula than chi-square, but the Fisher test has an important advantage over the chi-square test in that it does not involve approximations; rather, it calculates exact probabilities whatever the sample size.

Given the test design here, the Fisher test was able to reveal significant interactions between groups and items. A conservative alpha level was set for this test by inflating
the overall experimental wide error to .20 and dividing this number by the total number of comparisons (36). This set the significance level to $\alpha < .006$.

The basic approach was to question whether choice depended on group in a significant way. Using the Fisher test, significant interactions were found for 18 of 36 items. In an additional 13 of the 36 items, there were observable differences between responses that approached, but did not reach, the conservative significance threshold set for this test. Responses to only 5 of the 36 items displayed little or no difference between all levels. In general, acceptance of null subjects surfaced early, but was overgeneralized until the more advanced levels. Acceptance of inversion and that-trace effects clearly surfaced later, but the results of the inversion items are somewhat unclear, as we shall see.

### 5.3.1 Null subject results

There were two types of items related to null subjects: those requiring that topic-connected null subjects be dropped and those requiring that subjects that are not topic-connected be retained. Since English and Spanish differ only in regards to the former condition, the majority of the items in this task tested for the dropping of null subjects, but the task also included some items that tested for the retention of overt subjects in nontopic contexts. English and Spanish do not differ in their grammatical requirements regarding these contexts, so any differences surfacing for this condition would be interesting.
The results of the items testing for the topic subjects are given in Table 5.3, which provides the raw percentages of respondents choices (selecting null subjects) on each item by proficiency group:

Table 5.3 Percentage of choice for null subject, by item and proficiency group

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.5938</td>
<td>0.6667</td>
<td>0.8571</td>
<td>1.0000</td>
<td>0.8000</td>
</tr>
<tr>
<td>12</td>
<td>0.7344</td>
<td>0.7059</td>
<td>0.9286</td>
<td>0.8333</td>
<td>1.0000</td>
</tr>
<tr>
<td>14</td>
<td>0.5156</td>
<td>0.7255</td>
<td>0.5000</td>
<td>0.9333</td>
<td>0.8000</td>
</tr>
<tr>
<td>16</td>
<td>0.6406</td>
<td>0.6078</td>
<td>0.8750</td>
<td>0.8333</td>
<td>0.9333</td>
</tr>
<tr>
<td>18</td>
<td>0.6094</td>
<td>0.5490</td>
<td>0.6786</td>
<td>1.0000</td>
<td>0.7000</td>
</tr>
<tr>
<td>20</td>
<td>0.4375</td>
<td>0.4706</td>
<td>0.7857</td>
<td>0.8333</td>
<td>0.9333</td>
</tr>
<tr>
<td>31</td>
<td>0.4531</td>
<td>0.4706</td>
<td>0.4464</td>
<td>0.6667</td>
<td>0.6667</td>
</tr>
<tr>
<td>34</td>
<td>0.6719</td>
<td>0.6471</td>
<td>0.8393</td>
<td>0.8333</td>
<td>0.7333</td>
</tr>
<tr>
<td>Avg.</td>
<td>.58</td>
<td>.60</td>
<td>.74</td>
<td>.86</td>
<td>.82</td>
</tr>
</tbody>
</table>

For example, for dialogue 20 (below), native speakers preferred the null subject choice 93% of the time, while beginners and intermediates registered this preference only 43% and 47%, respectively.

(20) En la universidad
Simón: Fue muy difícil esa clase de biología.
Adriana: Ah, ¿sí? ¿Qué estudias?
Simón: Historia. No me gustan las ciencias.
Adriana: Qué lástima. Las ciencias pueden ser muy interesantes.
Simón: ¿Cómo te interesaste por las ciencias?
Adriana A: Tuve una maestra bienísima en la escuela secundaria. √
Adriana B: Yo tuve una maestra bienísima en la escuela secundaria.

The level of difference between native speakers and both beginners and intermediates for dialogue 20 is highly significant. Fishers Exact Test reveals a significance level of $p = .000001$ between advanced and beginners and $p = .000002$ between advanced and intermediates. Similarly, item 12 (below) registered differences between the beginning and intermediate groups compared to the native speakers. Beginners differed from native
speakers by selecting a null subject only 73% of the time, while natives always selected
the null choice. Intermediates also judged this sentences differently from natives,
selecting null subjects only 70% of the time.

(12) Pablo y Rosa en casa
Rosa: Pablo, ¿has visto mi dinero?
Pablo: ¿Dónde lo dejaste?
Rosa: Estaba aquí en la mesa.
Pablo: Mira debajo de esos documentos.
Rosa: ¡Ah, lo encontré!
Pablo A: Debes cuidar bien tu dinero. √
Pablo B: Tú debes cuidar bien tu dinero.

The differences between the beginning and intermediate groups and native speakers
registered significance levels of $p = .001$ and $p = .00006$, respectively.

The results on these items support a general pattern found also in other items (10,
16, 34) where beginners and intermediates fail to delete topic subjects at a rate similar to
the other three proficiency levels. This is generally true for the remaining topic-subject
items as well, since beginners and advanced learners in this study never deleted topic
subjects as often as native speakers, but the level of difference for these remaining items
fails to achieve significance.

Three items (17, 28, and 35) were predicted to require the retention of overt
subjects. For two of these items (17 and 28), the predicted result was obtained. On the
third item (35), native speakers did not perform as predicted, choosing the overt subject
only 40% of the time. Table 5.4 gives the percentages of choice for overt subject on each
item by proficiency group:
Table 5.4 Percentage of choice for overt subject, by item and proficiency group

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>0.6406</td>
<td>0.5882</td>
<td>0.6786</td>
<td>0.6667</td>
<td>0.8333</td>
</tr>
<tr>
<td>28</td>
<td>0.5781</td>
<td>0.4902</td>
<td>0.6607</td>
<td>1.0000</td>
<td>0.9333</td>
</tr>
<tr>
<td>35</td>
<td>0.1563</td>
<td>0.3333</td>
<td>0.6250</td>
<td>0.8333</td>
<td>0.4000</td>
</tr>
<tr>
<td>Avg.</td>
<td>.45</td>
<td>.47</td>
<td>.65</td>
<td>.83</td>
<td>.72</td>
</tr>
</tbody>
</table>

Items 17 and 28 (below) both supplied contexts in which contrastive or emphatic use of subject pronouns triggers overt realization. Native speakers chose overt subjects for these contexts at rates of 83% and 93%, respectively. In contrast, beginners selected overt subjects less frequently, 64% of the time for item 17 and 58% of the time for item 28. The difference from native speakers does not reach significance for item 17 ($p = .0891$), but does for 28 ($p = .0003$) (cp. with the results of the pilot on pp.155-156).

(17) Después de un paseo por el museo
Luis: Fui al museo esta mañana.
Rosa: ¿Y qué viste?
Luis: Muchas pinturas de Picasso.
Rosa: A mí me gusta mucho Picasso.
Luis: A mí también; y por eso compré un póster en la tienda de regalos.
Rosa A: Yo también compré uno la semana pasada. √
Rosa B: También compré uno la semana pasada.

(28) Haciendo una acusación
Carmen: ¿Entonces, qué está diciendo?
Rosa: Juana y usted estaban ahí en ese momento.
Carmen: ¿y?
Rosa: Que Juana y usted tenían una razón para hacerlo.
Carmen: Otro lo hizo.
Rosa B: Lo hicieron. Nadie más.

Of interest here is the observation that intermediates select overt subjects for items 17 and 28 at an even lower rate than beginners, and these rates achieve higher levels of significance for both items ($p = .0275$ and $p = .00003$, respectively). It should
also be noted that even the advanced group does not perform significantly better on these items, and near-natives patterned more with beginners, intermediates, and advanced than with natives on item 17.

The variable result among native speakers for item 35 (below) most likely arises from the unique discourse setting. The sentence preceding the binary choice in 35 appears to introduce a new discourse topic (car wrecks), which should require an overt subject to return to an earlier topic; however, for a majority of native speakers the salient topic, despite intervening material, appears to remain the two friends.

(35) *Malas noticias*

*Luis:* ¿Qué dijo tu hermano?  
*Julio:* No hay nada más que los doctores puedan hacer.  
*Luis:* ¡Qué horrible!  
*Julio:* Sí, es triste.  
*Luis:* Los choques de automóvil son muy trágicos.  
*Julio A:* Eran muy buenos amigos.  
*Julio B:* Esos dos eran muy buenos amigos.

Perhaps the saliency of a tragic death provides a persistent topic that here overrides typical topic behavior. Nevertheless, given this discourse situation, native speakers still retain overt subjects more frequently than beginners (40% to 15%, respectively), with a significance level of \( p = .0171 \).

The results on all the items related to null subjects may be summarized as follows. Learners begin to accept null subjects early, although they do not choose them as often as native speakers. It is not until the advanced level that choice patterns related to the dropping of topic-connect subjects converge on native patterns. Early learners also differ from natives in their choices regarding nontopic subjects. This is an unusual result because the choices made would not be preferred in English or Spanish, indicating that
learners create an interlanguage stage differing from both the L1 and L2. The difference between learners and native speakers becomes even more pronounced at the intermediate stage, where even fewer nontopic subjects are retained. Eventually, learners converge on the native patterns, but this happens only very late in the learning process.

**Section 5.3.2 Inversion results**

We turn next to the results related to inversion. The 12 inversion items were also of two types, six items that involved presentational or contrastive focus (for which it was predicted that inversion would occur) and six items that lacked this special discoursal feature (for which it was predicted that the uninverted choice would be preferred.) The investigation of whether inversion would occur with focused elements stemmed from Grimshaw and Samek-Lodovici (1995) and Samek-Lodovici (1996) who, as discussed in Section 4.4, argued that subjects in Italian obligatorily appear at the right edge of a VP when the discourse context calls for them to be focused. The presence of a pro-drop ‘parameter’ would imply a similar result for Spanish, but the results here suggest this is not the case.

Table 5.5 shows the percentages by group of how often subjects selected the inversion choice when no special focus was associated with that item.

**Table 5.5 Percentage of choices for inversion, without focus**

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.0938</td>
<td>0.0392</td>
<td>0.0714</td>
<td>0.5000</td>
<td>0.0333</td>
</tr>
<tr>
<td>3</td>
<td>0.1250</td>
<td>0.0196</td>
<td>0.0536</td>
<td>0.1667</td>
<td>0.1333</td>
</tr>
<tr>
<td>22</td>
<td>0.2344</td>
<td>0.1765</td>
<td>0.1250</td>
<td>0.1667</td>
<td>0.0333</td>
</tr>
<tr>
<td>23</td>
<td>0.1094</td>
<td>0.1961</td>
<td>0.2321</td>
<td>0.1667</td>
<td>0.0667</td>
</tr>
<tr>
<td>24</td>
<td>0.1406</td>
<td>0.0588</td>
<td>0.1964</td>
<td>0.1667</td>
<td>0.0000</td>
</tr>
<tr>
<td>30</td>
<td>0.2500</td>
<td>0.1373</td>
<td>0.1607</td>
<td>0.1667</td>
<td>0.0333</td>
</tr>
<tr>
<td>Avg.</td>
<td>0.16</td>
<td>0.10</td>
<td>0.14</td>
<td>0.22</td>
<td>0.05</td>
</tr>
</tbody>
</table>
All proficiency groups preferred not to invert for these items, as was expected. The Fisher Exact Test revealed that none of the differences in Table 5.5 were significant.

The results with respect to focused constituents were less clear. Table 5.6 displays the percentages of choices for inversion across proficiency levels for those items for which contrastive or presentational focus may have been thought to trigger inversion.

Table 5.6 Percentage of choices for inversion, with focus

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.2500</td>
<td>0.2157</td>
<td>0.3036</td>
<td>0.8333</td>
<td>0.9667</td>
</tr>
<tr>
<td>5</td>
<td>0.2344</td>
<td>0.1765</td>
<td>0.5357</td>
<td>0.6667</td>
<td>0.4667</td>
</tr>
<tr>
<td>8</td>
<td>0.3750</td>
<td>0.1765</td>
<td>0.2857</td>
<td>0.1667</td>
<td>0.2333</td>
</tr>
<tr>
<td>11</td>
<td>0.3125</td>
<td>0.2549</td>
<td>0.3393</td>
<td>0.3333</td>
<td>0.2667</td>
</tr>
<tr>
<td>15</td>
<td>0.2188</td>
<td>0.0392</td>
<td>0.1607</td>
<td>0.6667</td>
<td>0.2000</td>
</tr>
<tr>
<td>33</td>
<td>0.0469</td>
<td>0.0784</td>
<td>0.0714</td>
<td>0.5000</td>
<td>0.0333</td>
</tr>
<tr>
<td>Avg.</td>
<td>.24</td>
<td>.15</td>
<td>.28</td>
<td>.53</td>
<td>.36</td>
</tr>
</tbody>
</table>

In contrast with Italian, native speakers of Spanish do not regard inversion of these focused subjects as obligatory, nor do they even demonstrate a preference for inversion, with the single exception of item (4).

Nevertheless, several points should be noted. First, among native speakers, the average percentage of choices selecting inversion for focused constituents (36%) is higher than for non-focused constituents (5%). This is true also of each of the other proficiency levels, as Table 5.7 shows:

Table 5.7 Average percentage of selection for inversion, by condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-focus</td>
<td>.16</td>
<td>.10</td>
<td>.14</td>
<td>.22</td>
<td>.05</td>
</tr>
<tr>
<td>Focus</td>
<td>.24</td>
<td>.15</td>
<td>.28</td>
<td>.53</td>
<td>.36</td>
</tr>
</tbody>
</table>
Therefore, rather than inversion being forced by the discourse condition of focus (as Grimshaw and Samek-Lodovici (1995) claimed for Italian), in Spanish the effect of focus appears to be optionality. In other words, inversion may indeed be a ‘free’ choice in Spanish, at least as the choice relates to declaratives sentences and non-\textit{wh}-questions, such as those under investigation in this study.\footnote{6}

Second, we should notice that pairwise comparison of groups did yield two significant differences. For item (4) (below), native speakers strongly preferred (97\%) the inverted choice over the non-inverted choice (3\%). Near-natives also showed a strong preference for inversion on this item, while all remaining groups preferred the non-inverted choice.

(4) \textit{Asumiendo responsabilidad}

\begin{quote}
Pablo: ¿De qué te ríes? 
Janet: Mamá y Papá van a llegar pronto. 
Pablo: ¡Ay! ¡La casa está muy sucia! 
Janet: Te dije: ‘No invites a tus amigos.’ 
Pablo: ¿Me ayudas a limpiar la casa? 
Janet A: ¡No! Tienes que limpiarla tú, no yo. √ 
Janet B: ¡No! Tienes que limpiarla, no yo. 
\end{quote}

The highest level of significance for item (4) was between intermediates and natives ($p = .0000000005$), and beginners and advanced also registered significant differences from the near-native and native groups; however the results on this item must be viewed cautiously due to an additional consideration. The choices for (4) contrast an inverted subject with a null subject. Given this emphatic context, native speakers found the inverted subject preferable to the null subject, but their ideal choice may have been an overt, non-inverted subject.
The other significant interaction revealed was between intermediates and advanced learners on item 5:

(5) *En la oficina*

*Esther:* *Henry se va de vacaciones.*
*Juana:* ¿De veras? *El nunca se va de vacaciones.*
*Esther:* *Tienes razón pero esta vez sí se va.*
*Juana:* ¿A dónde va?
*Esther:* A Maui, pero no va solo.
*Juana A:* ¿Marta va con él? √
*Juana B:* ¿Va Marta con él? √

Item (5) did not have the confounding factor of item (4); it offered a true choice of inversion versus non-inversion. Native speakers regarded the choice of inversion here as optional, selecting the inverted choice 43% of the time. Advanced selected the inverted choice 53% of the time, and this was significant (*p* = .0001) in comparison with intermediates, who selected the inversion item only 17% of the time. In fact, it is only at the advanced level and above that choices for inversion ever reach more than 38% for any item.

Finally, although the near-native group converged on the native speaker group for an item such as dialogue 4 (where inversion was predicted), the near-native group generally inverted more than native speakers, whether the dialogue suggested focus or not. Near-native speakers inverted 22% of the time with non-focus items and 53% of the time with focus items—more than any other group. This suggests either hypercorrection, or that these speakers may enter a stage of overgeneralization, such as we saw with the intermediates regarding null subjects.

To summarize the findings regarding inversion, inversion in Spanish does not appear to pattern the same as has been claimed for Italian. Discourse does appear to have
an effect on the acceptability of inversion, but this effect is not an obligatory one. Recognition of the optionality of inversion for focused constituents appears most noticeably at the advanced level, with the lower levels selecting inverted choices an average of less than 25% of the time. The results here demonstrate convergence between natives and near-natives for some items, but this result is confounded by the possible interference of a choice between a null subject or an overt inverted subject in one instance, what appears to be true optionality in another other instance, and the possibility of overgeneralization on the part of the near-native group.

Section 5.3.3 *That-trace results*

The majority of the *that-trace* items in this study contrasted sentences with a complementizer *que* (followed by a null subject) with sentences lacking the complementizer (also followed by a null subject). A few additional sentences tested other manipulations of *que* and subjects: one sentence in which (+que/+null subject) was contrasted with (-que/-null subject), and two sentences in which (+que/-null subject) was contrasted with (+que/+null subject).

Those sentences contrasting a complementizer and null subject with sentences lacking the complementizer revealed clear and significant results. For all items, native and near-native speakers registered strong preferences for the inclusion of the complementizer on an average of 91% and 86%, respectively. This contrasted significantly with beginners, who registered a preference for complementizer inclusion only 46% of the time, and intermediates who did only slightly better (56%). The
advanced group performed squarely between the lower and higher levels, selecting the
complementizer 70% of the time.

Table 5.8 records the results by group and item in terms of the percentage of
choices for inclusion of the complementizer:

Table 5.8: Percentage of choices including complementizer (+que/+null vs. –que/+null)

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6563</td>
<td>0.6863</td>
<td>0.6250</td>
<td>1.0000</td>
<td>0.8667</td>
</tr>
<tr>
<td>6</td>
<td>0.4219</td>
<td>0.4118</td>
<td>0.5536</td>
<td>0.6667</td>
<td>0.8667</td>
</tr>
<tr>
<td>7</td>
<td>0.4062</td>
<td>0.3922</td>
<td>0.7143</td>
<td>0.8333</td>
<td>0.9333</td>
</tr>
<tr>
<td>9</td>
<td>0.4219</td>
<td>0.6275</td>
<td>0.6786</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>13</td>
<td>0.5469</td>
<td>0.8039</td>
<td>0.9821</td>
<td>1.0000</td>
<td>0.9667</td>
</tr>
<tr>
<td>19</td>
<td>0.4063</td>
<td>0.5294</td>
<td>0.7679</td>
<td>1.0000</td>
<td>0.8333</td>
</tr>
<tr>
<td>21</td>
<td>0.5469</td>
<td>0.5686</td>
<td>0.6071</td>
<td>0.5000</td>
<td>0.8667</td>
</tr>
<tr>
<td>25</td>
<td>0.4844</td>
<td>0.6078</td>
<td>0.6250</td>
<td>0.8333</td>
<td>0.9667</td>
</tr>
<tr>
<td>26</td>
<td>0.3437</td>
<td>0.4706</td>
<td>0.8571</td>
<td>1.0000</td>
<td>0.9667</td>
</tr>
<tr>
<td>27</td>
<td>0.4688</td>
<td>0.5294</td>
<td>0.7321</td>
<td>1.0000</td>
<td>0.9333</td>
</tr>
<tr>
<td>29</td>
<td>0.4375</td>
<td>0.5686</td>
<td>0.6964</td>
<td>0.6676</td>
<td>0.8333</td>
</tr>
<tr>
<td>32</td>
<td>0.5156</td>
<td>0.5882</td>
<td>0.6786</td>
<td>0.8333</td>
<td>0.9667</td>
</tr>
<tr>
<td>Avg.</td>
<td>.46</td>
<td>.56</td>
<td>.70</td>
<td>.86</td>
<td>.91</td>
</tr>
</tbody>
</table>

Item 9 (below) serves as an example of the pattern displayed regarding the
+que/+null vs. –que/+null sentences:

(9) En la tienda
Sr. Papas: ¿Puedo atenderlo?
Sr. Campos: Estoy buscando unos zapatos de tenis.
Sr. Papas: Estos son los mejores.
Sr. Campos: Pero son muy caros.
Sr. Papas: Sí, pero son de la más alta calidad.
Sr. Campos A: ¿Quién crees que los va a comprar a ese precio? √
Sr. Campos B: ¿Quién crees los va a comprar a ese precio?

For this type of item, the 2sg present form of creer ‘to think/believe’ requires the use of
que, and this is unfailingly recognized by the native and near-native speakers in this
sample. The Fisher Exact Test reveals highly significant differences between this native
speaker norm and the responses of beginning, intermediate, and advanced learners (\( p = .000000006, p = .00005, p = .0001 \), respectively). Therefore, native-like competence regarding this type of *that-trace* item is achieved here, but only at the near-native level.

A second type of *that-trace* interaction is represented by item 36 (below) in which a choice was given between +*que*/+null subject and –*que*/-null subject:

(36) *Haciendo planes para el fin de semana*

Laura: Estudias demasiado, Felicia.

Felicia: Tengo un examen en la clase de filosofía.

Laura: Estudiaste toda la semana y el fin de semana pasado.

Felicia: Voy a tomar un descanso pronto.

Laura: ¿Vamos a la playa mañana?

Felicia A: No creo yo pueda ir.

Felicia B: Yo no creo que pueda ir. √

Once again, native and near-native speakers registered a strong preference for the inclusion of *que* (86-100% of the time), but on this item, beginning through advanced learners also preferred the choice with *que* (70-80% of the time), and the level of difference between groups was not significant for this item.

This is an unexpected result for two reasons: first, the choice of including *que* in item 36 involves a violation of *that-trace*, and it might be expected that such a violation would not be the preferred choice in the grammar of a beginning learner of Spanish. Second, on the surface, this item appears to supply counterevidence to the implicational hierarchy of Liceras (1989), because it means that learners in this study first began choosing *that-trace* at the beginning and intermediate levels, while they did not chose items with inversion until the advanced level.

There is good reason to treat this conclusion with caution, however, because early learners may have made their choice on this item not because of the presence or absence
of *que*, but due to another difference between the items. The ‘A’ choice in item 36 has a null subject in the main clause, whereas the ‘B’ choice uses an overt subject. If early learners made their choice based on the overt subject, then we are left with the possible conclusion that both beginning and native groups made the same choice, but for different reasons.

The pattern apparent for item 36 also appears with items 14 and 18, items in which (+*que/-null subject) was contrasted with (+*que/+null subject). Item 18 is provided below as an example of this type:

(18) *Carmen y Felipe en la oficina*

*Carmen:* ¿Alguien me llamó cuando yo no estaba aquí?
*Esther:* Sí. Luis Pérez y Lilia Enríquez.
*Carmen:* Ok. ¿Alguien más?
*Esther:* No, no llamó nadie más.
*Carmen:* ¿Dijeron Luis y Lilia lo que querían?
*Esther A:* No, pero dijeron que iban a volver a llamar.
*Esther B:* No, pero dijeron que iban a volver a llamar. √

There were no significant differences between groups on items such as 18. In addition to testing whether a *that-trace* sequence is allowed, these items also provide an interaction between null and overt subjects. The percentages of choices for inclusion of the complementizer are given in Table 5.9.

Table 5.9: Percentage of choices including complementizer (+*que/+null vs. +*que/-null)

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Intermediate</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>0.5156</td>
<td>0.7255</td>
<td>0.5000</td>
<td>0.8333</td>
<td>0.8000</td>
</tr>
<tr>
<td>18</td>
<td>0.6094</td>
<td>0.5490</td>
<td>0.6786</td>
<td>1.0000</td>
<td>0.7000</td>
</tr>
<tr>
<td>Ave.</td>
<td>.56</td>
<td>.63</td>
<td>.58</td>
<td>.91</td>
<td>.75</td>
</tr>
</tbody>
</table>
It is tempting to conclude on the basis of items such as these that it is not the *that-trace* sequence that causes difficulties for the early learners since, for these items, learners sometimes choose *that-trace* over the option with an overt subject. Once again, some caution is required here, because the choice could be the result of the overall preference of respondents for null subjects. If this were the case, however, one would need to explain why early learners accepted the *that-trace* sequence in (36), an item with an overt subject in the matrix clause, and also accepted the *that-trace* sequence in items (14) and (28), items with null subjects. If learners are accepting *that-trace* even when there is an option that avoids it, then it is likely, or at least possible, that *that-trace* is not the source of the difference between their response patterns and those of native speakers. Future research on this issue will require more items of this type to provide stricter controls on conditions; results regarding this type of item in the present study remain somewhat tentative.

To summarize the findings on *that-trace*, when early learners were given a choice between sentences containing *that-trace* and sentences with a null subject but no complementizer, the learners included the complementizer a little less than 50% of the time. This contrasts sharply with native speakers, who consistently chose inclusion of the complementizer. When forced to choose between two sentences that include the complementizer but vary in regards to null or overt subjects in the subordinate clause, early learners demonstrated a slightly stronger preference for *that-trace* sequences, though at a rate significantly less than native speakers.

This result introduces the possibility that learners do not reject *that-trace* sentences based on the interaction between *que* and null subjects; rather, there may other
factors in play. For example, learners may not realize that the use of the complementizer in these contexts, which is frequently optional in English, is obligatory in Spanish. If so, then respondents are having less difficulty with the violation of syntactic constraint (e.g. *that-trace) than they are with identifying the conditions which require the use of que.

5.4 Interpretation

The results shown in this chapter now permit certain conclusions to be drawn regarding three of the four hypotheses presented in Section 4.1 of the previous chapter. The first of these hypotheses was that the results would indeed find evidence for the implicational hierarchy of Liceras (1989) (i.e. null subjects would appear before inversion, and inversion would appear before that-trace in the grammars of learners). The results of the study did not find strong support for the implicational hierarchy, due to the fact that, in this study, some that-trace effects appeared as early as the beginning or intermediate levels, and noticeably by the advanced level.

Nevertheless, the evidence here may have an interpretation within Liceras’ hierarchy for two reasons: First, although early learners did not often select the choice with inversion, the discourse condition of focus greatly raised the acceptance rate. Beginners accepted inversion 24% of the time with focus and only 16% without, and advanced learners accepted inversion 28% of the time with focus and only 14% without. In contrast, beginners double their average acceptance rate of that-trace only much later, at the advanced and near-native stages. Second, since inversion appears to be optional even for native speakers, little can be made of the fact that early learners use their L1 pattern when it does not conflict with the L2. The important question, then, is not why
learners choose inversion at a rate lower than that of native speakers, but what in their grammars permits them to choose inversion at all. The choice of inversion at a 24% rate, even though such a choice is optional in the target grammar, may reveal more than the acceptance of that-trace at a 40% rate when that choice is obligatory in the target grammar. For these reasons, Liceras’ hierarchy of null subjects > inversion > that-trace cannot be considered disconfirmed by this study’s results.

A second hypothesis that was presented stated that ‘initial acceptability’ would be distinct from ‘correct use’ for each of these grammatical properties. This hypothesis suggested that Liceras’s implicational hierarchy requires refinement, because initial acceptance of inversion does not imply a fully native grammar regarding null subjects, nor does initial acceptance of that-trace imply accurate use of inversion. There is enough evidence in this study to claim strong support for this hypothesis.

Learners in this study accepted null subjects very early, but their judgments regarding the acceptable use of null subjects did not demonstrate sensitivity to discourse conditions (topic/nontopic) until the advanced and near-native stages. Some time prior to learners’ correct usage of null subjects, that-trace begins to surface (though variably) in learners’ grammars. In the discourse context of focus constituents, choices for inversion increase among advanced and near-native groups. But near-native choices do not replicate native judgments regarding inversion as closely as for the null subject and that-trace conditions. All these findings demonstrate that a distinction must be drawn between initial acceptability (resulting from awareness of a particular grammatical property) and correct use (resulting from the proper ranking of discoursal constraints in regard to this grammatical property).
A third hypothesis presented claimed that at least some L2 learners would converge on native-like usage of null subjects, inversion and *that-trace*, and that this convergence would come as a result of sensitivity to discoursal constraints of the target language. The results of this study showed no statistically significant differences between near-native and native speakers. This was most clearly seen with *that-trace* and null subjects, where near-natives strongly converged on native response patterns. The results regarding inversion are less clear, largely due to the finding that inversion does not appear obligatory under any context even for native speakers. Near-native speakers do converge with native speakers even here, if ‘convergence’ is defined as ‘using inversion more frequently, but not obligatorily, with focused constituents’; but under such a broad definition, all groups (even beginners) can be said to converge on native speaker competence. Nevertheless, the question of ultimate attainment appears to have an affirmative answer here: the grammatical properties of null subjects, inversion, and *that-trace* can be acquired, by at least some learners, to a level such that L2 and native speakers do not differ significantly in their judgments.

These interpretations regarding an implicational hierarchy, initial acceptability vs. correct use, and ultimate attainment are summarized in Table 5.10:
Table 5.10: L2 developmental order: null subjects, inversion, *that-trace*

<table>
<thead>
<tr>
<th></th>
<th>Beginners</th>
<th>Intermediates</th>
<th>Advanced</th>
<th>Near-Native</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>null-subjects</td>
<td>selected at a lower rate than native speakers</td>
<td>selected at a higher rate than native speakers</td>
<td>appear to be selected with sensitivity to the discourse context, similar to native use</td>
<td>select when topics, retain when nontopics</td>
<td></td>
</tr>
<tr>
<td>inversion</td>
<td>selected at a much lower rate than native speakers, but rate increases when items are focused</td>
<td>inverted order is selected more frequently, but not obligatorily, in certain discourse contexts. Only weak evidence that learners converge on native judgments, some overgeneralization</td>
<td>optionally select inversion, but only when focused</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>that-trace</em></td>
<td>complementizer + null subject selected for certain contexts, but choices differ significantly from native speakers regarding the distribution of <em>que</em></td>
<td>learners start to select <em>that-trace</em> with greater frequency</td>
<td>selection of <em>that-trace</em> indistinguishable from native judgments</td>
<td>consistently select <em>that-trace</em> sequence</td>
<td></td>
</tr>
</tbody>
</table>

5.5 Limitations

The investigation presented here attempted to improve on previous studies in several ways: by asking subjects to make grammaticality judgments in the context of certain discourse conditions, by supplying choices intended to vary minimally in regards to the condition under investigation, and by supporting the findings with a production measure that would provide an additional measure of learner competence. All of these measures were taken, but several limitations remain.

First, although discourse contexts in the form of a dialogue were provided, certain interpretational differences regarding the saliency of topic or focus lingered. For some items it was unclear whether the reader was cueing off of the immediate topic of the preceding sentence or a more distant topic a couple of sentences earlier. Controlling for
this factor is difficult, because if the topic is held consistent throughout the dialogue, an ‘out-of-the-blue’ introduction of a new topic in the final response would seem very pragmatically odd. Nevertheless, the results of such an approach would be of interest.

Second, although the attempt was made to isolate the grammatical conditions under investigation through the use of sentential minimal pairs, even this approach permitted some seepage of conditions. For example, it is useful to investigate directly learners’ differences in judging between [+que/-null] and [+que/+null] sentences. One difference between such sentences is the presence or absence of that-trace, but these sentences also vary in the more simple sense of the null/non-null condition. Learners may make their selection based on this difference alone. One way that future research might begin to address this is first by testing a matrix sentence for the null/non-null distinction, and then by using the same sentence in a subordinate environment. That this approach was not taken here is a limitation of the present study.

Third, in terms of the distribution of subconditions (topic/nontopic, inverted/uninverted, etc.), the present study could have used more items that demonstrate clear native speaker preferences for inversion or retention of topic. The lower number of these items in the current study was the result of using the intuitions of a small group of native speakers for creating the test and conducting the pilot study using only intermediate learners. Intermediate learners were selected for the pilot because it was thought that this group would best reveal the approximate amount of time needed to administer the test and any difficulties with vocabulary; however, the results demonstrate that a more extensive pilot involving native speakers would have been useful.
A final limitation of the present study was that the translation task involved only a small number of total speakers from each level for a particular item. This was due to each subject’s translating only two dialogues. The problem with this approach is that, although it permitted conclusions to be made regarding significant differences between groups, it did not make specific item analysis possible. The reason each subject was not asked to translate every item was due solely to time and fatigue considerations. It would be a tiresome task to ask advanced speakers to translate all 36 dialogues, 216 lines of text, and it would far more daunting for beginners. One way around this problem for future studies might be to ask speakers to translate only the last sentence of each dialogue. This revised task would still likely involve a longer testing period than the current study, but it would be completable, particularly if a lower number of items for each condition (e.g. six, rather than 12) were given.
Notes

1 For this particular test, each item was answered by about only 8 people, so item analysis failed. This was not a problem in the larger grammaticality judgment task, in which every subject answered every item. It would have been tempting to simply create a table and look at the percentage of people that translated each item in a dispreferred way. The problem with this approach is that different groups of people responded to different items. A particular item (translated by only two representatives of a proficiency level) may appear to be ‘easy’ because it received a greater preponderance of preferred responses than it would typically, whereas if a different item received many dispreferred responses, then it appears that the latter item is ‘hard’, but this is confounded with the group of subjects that analyzed it. If this were easily modeled, there may have been a chance to disentangle group and item effects and arrive at adjusted item estimates. As it was, only group conditions were successfully modeled.

2 This was not an unexpected result. LaFond, Hayes, and Bhatt (2001) also showed that intermediate learners overgenerate null subjects in contexts where retention of the subject is required by the native grammar. Possible reasons for this overgeneration are suggested in Sections 6.2 and 6.3 of Chapter 6.

3 In the grammaticality judgment task, 83% of native speakers chose ‘A’ for (17) and 93% choose ‘A’ for (28). These two items did not undergo significant revision from the pilot to the grammaticality judgment test.

4 Translations of all dialogues in this chapter may be found in Appendix A.

5 After the death of a loved one, it would not be necessarily unusual, for example, for one sibling to tell another, ‘I can’t believe she’s gone’, even when the referent is not a topic in the conversation or immediate context.

6 This does not imply that inversion is never obligatory. For example, Baauw (1998) cites examples of obligatory inversion in wh-questions, both when the moved element is an argument (a) and when it is an adjunct (b), although inversion is sometimes optional in wh-adjuncts such as por qué (c) and cómo (d).

(a) ¿Qué dijo Juan ayer?
   *¿Qué Juan dijo ayer?
   ‘What (did) John say yesterday?’

(b) ¿Dónde escondió Juan el paraguas?
   *¿Dónde Juan escondió el paraguas?
   ‘Where (did) Juan hide the umbrella?’
Inversion with *wh*-elements does not involve the same type of focus-related constraint that is of interest in the current study. The relationship between *wh*-inversion and inversion based on focus is left for further research.
Chapter 6
Analysis and Conclusions

6.0 Introduction

The last chapter concluded with a discussion of this dissertation’s first three hypotheses in light of the results obtained here. The fourth, and final, hypothesis is repeated again in H4:

(H4) The developmental path taken by L2 learners may be characterized as the interaction of discrete discursal and syntactic constraints, and this interaction will confirm the operation of the Constraint Demotion Algorithm of Tesar and Smolensky (2000).

This chapter argues that this hypothesis is confirmed by the results of this study and presents the analysis used to support this claim.

The results of this study show that learners begin to drop subjects indiscriminately in the early stages of acquisition. At the next stage, learners appear to recognize that the target grammar requires the dropping of topics, not subjects. As sensitivity to discoursal conditions increases, learners’ selection of inversion also increases. Although early learners do not choose inversion as frequently as near-native and native speakers, even among these early learners the choice to invert is made more frequently when the discourse conditions call for the focusing of the subject. Unlike Italian, Spanish inversion in focused cases appears to involve a greater degree of optionality, in contrast to non-focused cases, where inversion simply does not occur. Although there is some
evidence that early learners may accept *that-trace* sequences, early learners only accept *that-trace* about half the time; native speakers consistently, obligatorily choose *that-trace* sequences. Learners eventually identify the discourse conditions related to the dropping of subjects, inversion, and *that-trace*. These are the facts for which an analysis of these grammatical properties must give an account.

In attempting to account for these facts, the analysis here will assume that the same universal constraints provided by UG are present in both English and Spanish, but that these constraints are ranked differently in these two languages. Acquisition occurs when learners detect the differences between the constraint rankings of the L1 and L2 and use input from the target language to recursively restructure their grammars to more closely match the grammar of the input. Since some aspects of the input may be more salient than others, not all differences may be detected simultaneously and acquisition may involve several rerankings. In the initial learning state, a fully hierarchized representation of the L1 is used to interpret target language data. Since the L1 constraint rankings quickly fail to adequately handle L2 input, Constraint Demotion is triggered and learning begins to take place.

Given this set of assumptions, the analysis in this chapter argues that the developmental stages in the acquisition of null subjects, inversion, and *that-trace* can be understood as an interaction between the syntactic and discoursal constraints given in Figures 6.1–6.2 (repeated from Figures 1.1–1.2 in Chapter 1):
Figure 6.1  Syntactic constraints

a.  **SUBJECT (SUBJ):** The highest A-specifier in an extended projection must be filled. Failed when a clause lacks a subject in the canonical position. (Grimshaw 1995)
b.  **PARSE:** Parse input constituents. Failed when input elements are not overtly parsed in the output.
c.  **FAITH[SUB]:** The output value of [SUB] (for ‘subordination’) must be the same as the input value. (Baković 1997)
d.  **T-LEX-GOV:** A trace is lexically governed. (Grimshaw 1997)

Figure 6.2  Discoursal constraints

a.  **ALIGNFOCUS-RIGHT (AF-RT):** Align the left edge of focus constituents with the right edge of a maximal projection. Failed by non-aligned foci. (Grimshaw and Samek-Lodovici 1995)
b.  **DROPTOPIC (DROPT):** Leave arguments coreferent with the topic structurally unrealized. Failed by overt constituents which are coreferential with the topic. (Grimshaw and Samek-Lodovici 1995)

The classification of constraints as ‘discoursal’ or ‘syntactic’ is based upon whether the constraint affects the realization of units within clauses, in which case it is considered syntactic, or whether it affects larger structures or ordering of information beyond the clause, in which case it is considered discoursal. **SUBJECT, PARSE, FAITH[SUB],** and **T-LEX-GOV** relate to the realization of constituents within a clause, determining the correct set of syntactic properties for a language; therefore, these constraints are here considered syntactic. Among these, **PARSE** and **FAITH[SUB]** belong to the same family of ‘faithfulness’ constraints, constraints that require the input and the output of the grammar to be identical.

In contrast, **ALIGNFOCUS-RIGHT** and **DROPTOPIC** clearly relate to the ordering or realization of information based on extra- clausesal information. Information packaging requirements are more fluid in a language, varying from utterance to utterance in
accordance with contextual considerations. Therefore, these constraints are considered discoursal.

The constraints in Figures 6.1–6.2 will be discussed further as they are used in the analysis, but the general observation is that the acquisition of Spanish by native speakers of English involves a dynamic re-ranking of constraints — most notably a demotion of certain syntactic constraints with respect to the discoursal constraints mentioned above.

Chapter 2 mentioned some possible types of alterations to a constraint hierarchy that might occur (from Hutton 1996). In terms of constraint demotion, those possibilities may be reduced to three: (1) demotion of one constraint (A) below another (B) where dominance relations are reversed ($A$$\Rightarrow$$B$ $\rightarrow$ $B$$\Rightarrow$$A$), (2) demotion of a constraint (B), previously tied or unranked in regard to another constraint, to a position where that constraint is now dominated ($\{A,B\}$ $\rightarrow$ $A$$\Rightarrow$$B$), or (3) demotion of a constraint (A), previously in a dominance relation in regard to another constraint, to a position where it is now tied with the constraint that formerly dominated it ($A$$\Rightarrow$$B$ $\rightarrow$ $\{A,B\}$). All three of these types of demotions are evidenced in the development of the grammatical properties analyzed here.

Section 6.1 begins with a more detailed description of how the mechanism of constraint demotion works in the special case of tied constraints ($\{A,B\}$) or stratified domination hierarchies ($\{A,B\}$$\Rightarrow$$\{C,D\}$), as is the case in some of the interactions here. Sections 6.2 proposes specific demotions involved in the realization of null subjects, both in the early stages of language acquisition, during which null subjects are overgeneralized, and in the later stages, in which learners discover that topics, not subjects, must be dropped. Section 6.3 considers the demotion process that moves
learners from a ranking where inversion is not possible, to one where it is possible but not obligatory. It is argued here that this demotion is connected to the acquisition of null subjects, suggesting one reason why null subjects and inversion often pattern together in pro-drop languages. Section 6.4 then takes up the issue of *that-trace* and provides two different analyses of the stages learners go through related to the complementizer: a first analysis which accounts well for the data, but involves uncertain modifications to the learning algorithm, and a second analysis which better honors the integrity of the learning algorithm but involves an expanded definition of input. Section 6.5 provides a brief summary of the analysis, pulling all the demotions involved in the acquisition of these properties together into one figure. Finally, Section 6.6 concludes the dissertation by summarizing what has been learned, to what degree this account of constraint demotion improves upon parameter-setting models, and how the findings suggest further lines of research in SLA.

### 6.1 Constraint demotion in stratified domination hierarchies

The normal operation of constraint demotion contrasts two constraints and changes the dominance relationship between them such that A»B becomes B»A; however, constraint demotion is also able to produce cases in which moving one constraint does not place it below another, creating the possibility of ties between constraints. This section will show how this result is obtained and why there is good reason to believe that certain constraints are either unranked in relation to each other (A,B), or if they are in a tie, they have two separate but equal rankings (i.e. A»B and B»A are equally valued).
Tesar and Smolensky (2000) provide for the possibility of constraint demotion in the case of nonranked constraints by arguing that their learning algorithm operates within the space of stratified hierarchies, represented in Figure 6.3:

Figure 6.3  Stratified domination hierarchy (Tesar and Smolensky 2000:37)

\{C_1, C_2, ...., C_3\} » \{C_4, C_5, ...., C_6\} » ... » \{C_7, C_8, ...., C_9\}

The idea here is that although constraints of each stratum are not ranked with respect to each other, each of the constraints \(C_1\), \(C_2\), and \(C_3\) dominate those of the next stratum and subsequent strata.\(^2\)

In this chapter, following standard practice, commas between constraints indicate this nonranking, and the » symbol indicates a dominance relationship. In the tableaux in this chapter, also following standard practice, nonranking between constraints will be indicated by dotted lines; whereas, constraints that are in a dominance relationship to one another will be indicated by solid lines. Cells to the right of fatal violations are shaded to show that those constraints are irrelevant to the evaluation of candidates, since the candidate has already been eliminated by the fatal violation. These conventions are illustrated in Tableau 6.1 (adapted from Tesar and Smolensky 2000:38):

Tableau 6.1  Stratified hierarchy \(C_3 \rightarrow \{C_4, C_5\} \rightarrow C_6\)

<table>
<thead>
<tr>
<th></th>
<th>(C_3)</th>
<th>(C_4)</th>
<th>(C_5)</th>
<th>(C_6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>candidate A</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>candidate B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>candidate C</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>candidate D</td>
<td></td>
<td></td>
<td></td>
<td>**!</td>
</tr>
</tbody>
</table>
The optimal candidate in Tableau 6.1 is candidate C. A single violation of either of Constraint 4 or Constraint 5 would not be sufficient in this tableau to disqualify the candidate from further consideration. Candidate A fails because it violates the highest ranking constraint here. Candidate B fails because it not only violates a constraint in the second stratum but also violates an additional constraint in the next stratum. Finally, candidate D fails because multiple violations within the same stratum are worse than a single violation.

Constraint demotion, which operates through a pairwise comparison of candidates, may, therefore, affect a grammar in three ways. First, a comparison between Constraint 3 and Constraint 4 above may demote Constraint 3 from the first stratum to a position below the second stratum. This yields the possibility of a simple reversal of the dominance relationship ($C_3 \rightarrow \{C_4, C_5\} \rightarrow \{C_4, C_5\} \rightarrow C_3$). Second, a comparison between Constraints 4 and 5 could require that Constraint 5 be demoted to the next stratum, where it is now dominated by Constraint 4, but on the same tier as Constraint 6. This yields the possibility of moving from a position of nonranking to a new dominance relationship that is shared with another constraint ($\{C_4, C_5\} \rightarrow C_6 \rightarrow C_4 \rightarrow \{C_5, C_6\}$). Finally, a comparison of Constraint 5 and Constraint 6 could require that Constraint 5 be ranked below Constraint 6. This is again a simple reversal of dominance, but differs from the earlier demotion (of $C_3$ below $C_4$) in that here there is a change in the original dominance relationship not only with respect to Constraint 5, but also with respect to Constraint 4 ($\{C_4, C_5\} \rightarrow C_6 \rightarrow C_4 \rightarrow C_6 \rightarrow C_5$). With the concept of stratified domination hierarchies in mind, we are now able to turn to the analysis of this chapter.
6.2 Null subjects

The analysis in this section begins with the assumption that beginning learners transfer their native language constraint hierarchy into their learning of the L2 (Schwartz and Sprouse 1996). It also assumes the initial constraint rankings for English motivated by Grimshaw and Samek-Lodovici (1995) and LaFond, Hayes, and Bhatt (2001) — that the English hierarchy is one in which a faithfulness constraint requiring the parsing of input constituents (PARSE) dominates both the syntactic constraint requiring clauses to have overt subjects (SUBJECT) and the discoursal constraint requiring the dropping of topic-connected subjects (DROP_TOPIC). This hierarchy will always require overt subjects, regardless of discourse condition. In contrast, the ranking of the target (Spanish) hierarchy must be one that produces overt subjects only when they are non-topics.

The observation that intermediate learners indiscriminately drop subjects, not just topics, suggests a stage of acquisition in which learners recognizes that Spanish places a lesser priority than English on the parsing of input, but does not yet recognize the extent to which this difference in priorities is determined by discourse conditions. This section argues that the developmental path learners take in their acquisition of null subjects is characterized by the rankings in Figure 6.4 (below).

Figure 6.4 Rankings yielding non-null/null subjects

a. English ranking: \text{PARSE} \gg \text{SUBJECT} \gg \text{DROP_TOPIC} \gg \text{ALIGN_FOCUS-RT}

b. Intermediate ranking: \text{SUBJECT} \gg \text{DROP_TOPIC} \gg \{\text{PARSE, ALIGN_FOCUS-RT}\}

c. Spanish ranking: \text{DROP_TOPIC} \gg \{\text{PARSE, SUBJECT, ALIGN_FOCUS-RT}\}
Although the rankings of PARSE, SUBJECT, and DROP_TOPIC in Figure 6.4a follow the analyses of Grimshaw and Samek-Lodovici (1995) and the developmental account of null subjects of LaFond, Hayes, and Bhatt (2001), the current analysis’ subsequent rerankings involve some differences from previous accounts. Whereas LaFond, Hayes, and Bhatt (2001) proposed that learners of Spanish first demote SUBJECT and later PARSE and that the final Spanish ranking is one in which PARSE still dominates SUBJECT, the analysis here demotes PARSE before SUBJECT and holds that the final Spanish ranking is one in which PARSE is on the same stratum as SUBJECT (and also ALIGN_FOCUS_RIGHT).

The proposal of LaFond, Hayes, and Bhatt was motivated by the premise of Grimshaw and Samek-Lodovici (1995) that PARSE must dominate SUBJECT because the null candidate in Tableau 6.2, which leaves all input unparsed, would satisfy the SUBJECT constraint.

<table>
<thead>
<tr>
<th>Input</th>
<th>DROP</th>
<th>PARSE</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ha cantato (‘has sung’)</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. Ø</td>
<td>**!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

But this premise is only true if the radically null candidate does not violate SUBJECT, and it remains unclear why this should be so. The input in Tableau 6.2 has a subject, lui, and the null candidate fails to parse this subject. The assumption being made is that the definition of SUBJECT is such that a clause without one is violated, but if there is no clause there is no violation (Jane Grimshaw, personal communication). If no input material is parsed at all, there is no violation of SUBJECT, because there will be no clausal structure.
Nevertheless, to hold that PARSE must dominate SUBJECT because of the radically null candidate requires the assumption that null candidates involve not only leaving structural positions unfilled but, also, the removal of that structure altogether. This is problematic for those who would maintain that GEN uses input to produce candidate structures in keeping with X-bar Theory. Fortunately, there is no need to assume that PARSE must dominate SUBJECT in Italian (or Spanish), since assuming that the null parse also violates SUBJECT produces identical results.

Tableau 6.3 PARSE and SUBJECT in Italian reconsidered

<table>
<thead>
<tr>
<th>Input &lt;cantare (x), x=topic, x=lui; T=pres perf&gt;</th>
<th>DROP</th>
<th>PARSE</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>a. ha cantato (‘has sung’)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>K</td>
<td>b. Ø</td>
<td>**</td>
<td>*(!)</td>
</tr>
</tbody>
</table>

In Tableau 6.3, the non-ranking of PARSE with regards to SUBJECT means that the radically null candidate will not be selected because it contains an additional violation not found in the winning candidate. Although the radically null parse is here shown with two violations (for the verb and its argument), another possibility is that leaving tense unparsed incurs a third violation, and this would also eliminate the null candidate, regardless of whether PARSE dominates SUBJECT or is positioned on the same tier as SUBJECT. If PARSE and SUBJECT are not ranked with respect to each other, any violations of either constraint that cause the total number of violations for one candidate to exceed another will be fatal. In the present example, two violations on the second tier in Tableau 6.3 are acceptable, but a third violation is fatal.

Although PARSE and SUBJECT do not appear to be ranked in Italian, there is some evidence that they are ranked in English in the order that LaFond, Hayes, and Bhatt
(2001) suggest. In English, PARSE and SUBJECT do not normally enter into competition with one another, but in those rare instances that they do, a violation of PARSE is worse than a violation of SUBJECT. For example, a sentence such as He said ‘I agree’, may undergo a fronting operation that moves the lower sentence to the head of the matrix sentence. Such fronting does not necessarily violate PARSE or SUBJECT (6.1a); however, when elements are fronted in this way, subject inversion may occur (6.1b), resulting in a violation of SUBJECT, but not PARSE. Although such sentences are involve special discoursal contexts, there is a clear contrast between violating SUBJECT, as 6.1b does, and violating PARSE, as in 6.1c-e.

(6.1.) a. ‘I agree’ he said.
   b. ‘I agree’ said he.
   c. ‘I agree’ said.
   d. *Ø he said.
   e. *He said Ø.

The difference between the candidates that violate SUBJECT and the candidate that violates PARSE is shown in Tableau 6.4:

Tableau 6.4 PARSE and SUBJECT in English

<table>
<thead>
<tr>
<th>Input</th>
<th>PARSE</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ‘I agree’ he said.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ‘I agree’ said he.</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. ‘I agree’ said.</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>d. Ø he said.</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>e. He said Ø.</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>
Candidates in 6.4a and 6.4b show a gradient ranking. The favored structure in English is 6.4a, and this response violates neither PARSE nor SUBJECT. Although the 6.4b ought to be ungrammatical according to the grammar, apparently violation of SUBJECT is not enough in this discourse style to eliminate the candidate. Hence, candidates that violate SUBJECT, are sometimes allowed in English, but violation of PARSE generally proves fatal. This result is different than Spanish, where the competition between PARSE and SUBJECT no longer distinguishes candidates due to the fact that in Spanish both PARSE and SUBJECT are dominated by the higher ranked constraint, DROPTOPIC. Based on these observations, the analysis in this dissertation assumes the same initial constraint ranking of PARSE and SUBJECT for English proposed by Grimshaw and Samek-Lodovici (1995) and LaFond, Hayes, and Bhatt (2001), but proposes that this constraint ranking is not preserved in Spanish, where PARSE and SUBJECT are positioned on the same stratum.

If PARSE dominates SUBJECT in English but not in Spanish, and if PARSE and SUBJECT may terminate on the same stratum, the question then becomes which demotes first as learners acquire Spanish. The study of LaFond, Hayes, and Bhatt (2001) could not, based on their study of only null subjects, demonstrate any empirical reason why SUBJECT should demote before PARSE. By looking at a broader picture that includes inversion, it becomes clear that PARSE must first demote, because the demotion of SUBJECT below DROPTOPIC places SUBJECT on the same tier as ALIGNFOCUS-RIGHT. As will become clear in the next section, the demotion of SUBJECT to the same tier as ALIGNFOCUS-RIGHT creates the inversion effects found in Spanish, and since these effects do not surface until after the acquisition of null subjects, we can establish the order of demotion — PARSE demotes first, followed by the demotion of SUBJECT.
Given the constraint rankings established in Figure 6.5, we are left with the question of whether constraint demotion effectively moves learners from one stage to the next. Application of the Constraint Demotion Algorithm (Tesar and Smolensky 2000) would operate in the following manner:

Given the initial English constraint hierarchy, a native English speaker chooses candidate (b) of Tableau 6.5a as the optimal output candidate when the subject pronoun is a topic, and also when the subject is a non-topic (6.5b):

Tableau 6.5 English Ranking: PARSE » SUBJ » DROP_T » AF-RT

<table>
<thead>
<tr>
<th>Subject is a topic</th>
<th>PARSE</th>
<th>SUBJ</th>
<th>DROP_T</th>
<th>AF-RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø has [Ø smiled]]</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [IP he_i has [t_i smiled]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. [IP Ø has [[t_i smiled] he_i]]</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject is not a topic</th>
<th>PARSE</th>
<th>SUBJ</th>
<th>DROP_T</th>
<th>AF-RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø has [Ø smiled]]</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [IP he_i has [t_i smiled]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. [IP Ø has [[t_i smiled] he_i]]</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When this speaker of English begins to learn Spanish, she quickly encounters sentences with null subjects, and her current English grammar makes the wrong choices with such sentences. Since the current grammar selects null subjects at a lower rate than the target grammar requires, the Constraint Demotion Algorithm (Tesar and Smolensky 2000) is activated, and the grammar begins to evaluate winner/loser pairs of candidates against the constraint hierarchy. Consider an input such as 6.2 where a topic subject is used in response to the question, ‘Where did Juan go?’
(6.2.) ∅ Fue a la playa.
go-3sg-pst to the beach
‘He went to the beach.’

As noted above, although both PARSE and SUBJECT must be demoted below DROP Topic, the developmental results help us establish that the first constraint demoted is PARSE.

Tableau 6.6 shows the first loser/winner pair:

Tableau 6.6 Mark-data pair: Beginning grammar (null subjects)

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>PARSE</th>
<th>DROP T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser: [IP él [VP ti fue a la playa]]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Winner: [IP Ø [VP Ø fue a la playa]]</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

In Tableau 6.6, the target winner is not optimal according to the current (loser) grammar, requiring the operation of the Constraint Demotion Algorithm.

The CDA looks for the highest-ranked of loser violations. Given the logic of the learning algorithm, each constraint in the winner candidate is checked to see if it is dominated by the loser marks. In the above example, PARSE is not dominated by the loser mark DROP Topic. Therefore, the CDA requires PARSE be demoted to the stratum immediately below that of the loser mark, DROP Topic (Tableau 6.7):

Tableau 6.7 Demotion of PARSE

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>PARSE</th>
<th>SUBJ</th>
<th>DROP T</th>
<th>AF-Rt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser: [IP él [VP ti fue a la playa]]</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Winner: [IP Ø [VP Ø fue a la playa]]</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The resulting ranking still has the SUBJECT constraint in the winner candidate dominating the loser mark DROP Topic, so the CDA will eventually need to go into
operation once again. Before that additional movement is shown, we should first consider the status of the learner after the demotion that has yielded Tableau 6.8:

Tableau 6.8: Ranking after the demotion of PARSE

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>SUBJ</th>
<th>DROPT</th>
<th>PARSE</th>
<th>AF-Rt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser:  [IP él, [VP ti fue a la playa]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winner:  [IP Ø [VP Ø fue a la playa]]</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Tableau 6.8 shows that a restructuring of the grammar has taken place that moves the learner closer to the target language, but the resulting grammar at this stage is neither that of English nor of Spanish. The demotion of PARSE appears to be a radical step for English speakers and leaves them with uncertainty about what may be dropped in the target language and what may not. It is at this stage that learners take the unwarranted and unexpected step of dropping all subjects, not just those that are topic-connected. The grammar in Tableau 6.8 does not predict this result, if we assume that learners’ outputs should be the result of their constraint rankings. It is clear that learners deviate from the learning their grammar has achieved, and this forces us beyond the formal account to recognize that interlanguage grammars are inherently unstable, both because interlanguage constraint hierarchies may fluctuate for a time and because learners may use general cognitive strategies to construct ad hoc hypotheses about the language they are learning.

LaFond, Hayes, and Bhatt (2001) attributed the simultaneous presence of target-like and non-target-like structures to a ‘grammar in flux’. They argue that interlanguage stages may move from domination, to tendency, to optionality on their way to a reversal of dominance relations (Figure 6.5).
Figure 6.5 shows that the interaction of two constraints, c and p, in second language acquisition may involve more than a simple demotion of one constraint below the other. Dominance relationships may be weakened to tendencies (represented by the > symbol) or even optionality, and rankings displaying optionality or tendency may strengthen into a dominance relationship. ‘Tendency’ here refers to a preference for a particular constraint ranking (perhaps because that ranking is more familiar or easily learned). This tendency is one which may, perhaps uncomfortably, be violated under certain conditions. Several factors may impinge upon this preference, including the interplay of tied or non-ranked constraints in stratified hierarchies and general learning strategies that do not instantiate UG-based options in the grammar. The differences between domination, tendency and optionality can be illustrated even without reference to learner grammars. For example, one tendency mentioned earlier in this chapter related to the grammar’s choice of ‘I’ll go’ John said vs. ‘I’ll go,’ said John. A grammar with a particular constraint ranking might ‘tend’ to choose the prior of those two sentences, since that sentence does not violate SUBJECT, but while the grammar may maintain a preference for this choice, it may also permit the latter option under certain conditions. If there is no clear preference between two choices, as is sometimes thought to be the case.

![Diagram of Grammar in Flux](image-url)

<table>
<thead>
<tr>
<th>L1</th>
<th>IL1</th>
<th>IL2</th>
<th>IL3</th>
<th>IL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>b » c » p » q</td>
<td>b » c &gt; p » q</td>
<td>b » c, p » q</td>
<td>b » p &gt; c » q</td>
<td>b » p » c » q</td>
</tr>
</tbody>
</table>

(Initial State)                                (Steady State)

- **optionality**
- **tendency**
- **domination**
in a pair of sentences such as *I said I’ll go* vs. *I said that I’ll go*, then we might understand the constraint rankings producing equally grammatical candidates as a case of true optionality.

If we consider learner grammars as works in progress, then the interplay of dominance, tendency, and optionality may be magnified. If Figure 6.5 (above) illustrates the functioning of a grammar within the unstable space of learner grammars, then we may better understand an environment characterized by hypothesis testing, backsliding, and failure as learner progressively move toward the target grammar. Further research into the possibility of grammars in flux may yield a better understanding of the relationship between UG-based hypotheses and general learning strategies, a concern that would help further explain the results obtained in this study.

The unstable character of learner grammars suggests one possible explanation of why learners may act in a manner not predicted by the state of their grammars. We may also note that the construct of ‘subject’ is far more salient than ‘topic’ in most language learning and instruction, and that learners are sometimes explicitly told that subjects are usually omitted in Spanish unless they are ambiguous or needed to indicate emphasis (Eric Holt, personal communication). Learners whose L1 make little use of topic status may need to learn what constitutes a topic (LaFond, Hayes, and Bhatt 2001), and it is understandable that the demotion of PARSE may leave learners uncertain as to which elements can be omitted and which can not. If the language instruction that learners have received has failed to distinguish between subject and topics, then we should not be surprised if learners resort to an ad hoc, non-UG-directed, rule in the face of a grammatical system that is in the process of restructuring but is not yet native-like.6
Returning to the learning algorithm, it is apparent that the ranking in Tableau 6.8 does not yet yield the target results (L2). Regardless of the reasons behind the overgeneralization of null subjects at the intermediate stage, learners eventually become sensitive to the fact that subjects in Spanish are dropped only when they are topics. If Juan in 6.3 (below) instantiates a non-topic, the target grammar requires the subject NP to be overt; if it instantiates a topic, that topic should be dropped.

(6.3.) \textit{Juan/∅ fue a la playa.}
Juan go-3sg-pst to the beach
‘Juan went to the beach.’

Since the interlanguage grammar of intermediates does not yet produce this result, further operation of the CDA is required. The grammar once again evaluates winner/loser pairs of candidates against the constraint hierarchy and now must demote the winner mark (SUBJECT) below the loser mark (DROP TOPIC), shown in Tableau 6.9:

Tableau 6.9  Demotion of SUBJECT

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>SUBJ</th>
<th>DROP'T</th>
<th>PARSE</th>
<th>AF-R T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser: [IP él [VP ti fue a la playa]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winner: [IP ∅ [VP ∅ fue a la playa]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Following the demotion of SUBJECT, the grammar (Tableau 6.10a and 6.10b) now properly drops subjects only when they are topics. All other subjects are retained.

Tableau 6.10a  Advanced grammar (subject is a topic)

<table>
<thead>
<tr>
<th></th>
<th>DROP'T</th>
<th>PARSE</th>
<th>AF-R T</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. K [IP ∅ [VP ∅ fue a la playa]]</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. [IP Juan [VP ti fue a la playa]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tableau 6.10b  Advanced grammar (subject is a non-topic)

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>DROP</th>
<th>PARSE</th>
<th>AF-RT</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø [VP Ø fue a la playa]]</td>
<td>*(!)</td>
<td>⋮</td>
<td>⋮</td>
<td>*</td>
</tr>
<tr>
<td>b. Κ [IP Juan [VP t₁ fue a la playa]]</td>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
<td>⋮</td>
</tr>
</tbody>
</table>

Since this ranking correctly produces target outputs, the learning algorithm halts in regards to null subjects. Learners have converged on the target grammar in a two-step process, first by demoting PARSE in relation to DROP TOPIC, and then by demoting SUBJECT in relation to DROP TOPIC. Subjects are now dropped when topic-connected and retained when they do not instantiate a topic. Constraint demotion is thus able to yield the changes in the grammar related to null subjects that are suggested by the results of this study.

6.3 Inversion

The constraints shown thus far permit a description of the developmental path learners take in the acquisition of null subjects, but they do more than that; they also show why null subjects and inversion often pattern together — the second of the constraint demotions that takes place in the learning of null subjects is a demotion that results in the allowance of inversion.

To understand this result, it must be observed that the demotion that places SUBJECT on the tier below DROP TOPIC, places that constraint on the same tier as ALIGN FOCUS-RIGHT. Prior to this demotion, the requirement that the highest A-specifier in an extended projection be filled meant that the effects of ALIGN FOCUS-RIGHT would not surface, due to it always being dominated by SUBJECT. Following this demotion, the grammar has the option of filling either SUBJECT or ALIGN FOCUS-RIGHT, even though
satisfying one of the constraints will involve a violation of the other, if a focused element is in the input.

ALIGNFOCUS-RIGHT is a member of a general class of alignment constraints that permits consistency in the information ordering of the sentence. The definition of this constraint given by Grimshaw and Samek-Lodovici (1995) (cf. Figure 6.2 on p. 183) includes a directionality of the focusing, aligning the left edge of focus constituents with the right edge of a maximal projection. This is also the definition that is used here. Just as English learners of Spanish need to demote certain syntactic constraints in their L1 grammar below the discoursal constraint of DROP TOPIC, so also must they recognize the need to maintain the level ranking of the discoursal constraint ALIGNFOCUS-RIGHT in regards to SUBJECT, despite the fact that in English ALIGNFOCUS-RIGHT is below SUBJECT. With respect to each other, SUBJECT and ALIGNFOCUS-RIGHT may be ranked in three different ways, and each of these ways results in a natural human language (Figure 6.6):

Figure 6.6  Ranking SUBJECT and ALIGNFOCUS-RIGHT

(a) SUBJECT » ALIGNFOCUS-RIGHT (English)
(b) ALIGNFOCUS-RIGHT » SUBJECT (Italian)
(c) SUBJECT, ALIGNFOCUS-RIGHT (Spanish)

Consider the different predictions these rankings make in Tableaux 6.11–6.13:

Tableau 6.11  SUBJECT » ALIGNFOCUS-RIGHT (English)

<table>
<thead>
<tr>
<th>A. &lt;wept (x), x=focus, x=he; T=pres perf&gt;</th>
<th>SUBJ</th>
<th>ALIGNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø has [ Ø wept]]</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>Kb. [IP he t has [ t wept]]</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. [IP Ø has [[ t wept he]]]</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Tableau 6.12  ALIGNFOCUS-RIGHT » SUBJECT (Italian)

<table>
<thead>
<tr>
<th>A. &lt;wept (x), x=focus, x=he; T=pres perf&gt;</th>
<th>ALIGNF</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø has [ Ø wept]]</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b. [IP he_i has [ t_i wept]]</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>K c. [IP Ø has [ [ t_i wept] he_i ]]</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 6.13  SUBJECT, ALIGNFOCUS-RIGHT (Spanish)

<table>
<thead>
<tr>
<th>A. &lt;wept (x), x=focus, x=he; T=pres perf&gt;</th>
<th>SUBJ : AF-Rt</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ø has [ Ø wept]]</td>
<td>*</td>
</tr>
<tr>
<td>K b. [IP he_i has [ t_i wept]]</td>
<td>*</td>
</tr>
<tr>
<td>K c. [IP Ø has [ [ t_i wept] he_i ]]</td>
<td>*</td>
</tr>
</tbody>
</table>

The tableau in 6.11, where SUBJECT dominates ALIGNFOCUS-RIGHT, will not permit inversion, while Tableau 6.12, where ALIGNFOCUS-RIGHT dominates SUBJECT, will require inversion of focus constituents. Tableau 6.13, in contrast to both of these, represents a grammar where the two constraints are unranked (or ‘tied’) with respect to each other.

To move from the ranking in 6.11 (English) to the ranking in 6.13 (Spanish), one might assume that learners must contrast SUBJECT directly with ALIGNFOCUS-RIGHT and realize that the dominance relationship must be undone, so that the constraints are now nonranked; however, CDA only demotes constraints to a stratum below a competing constraint. Given this operation of the algorithm, moving from a dominance relationship to a nonranked position must involve an interaction of two constraints at a higher stratum, not a direct interaction between the constraints of the higher stratum and the stratum to which that constraint is moving. Such a competition between two higher-ranking constraints does exist, namely the competition between SUBJECT and DROPTOPIC. Until learners demote SUBJECT below DROPTOPIC, SUBJECT is positioned in a dominance relationship to ALIGNFOCUS-RIGHT. It is only this demotion that positions SUBJECT and
ALIGNFOCUS-RIGHT in a tied-relationship and creates the optionality of inversion that Spanish exhibits.

The acceptance of inversion, therefore, is an epiphenomenon, a by-product of a grammar that either ranks SUBJECT above or below DROPTOPIC. To see this, consider again the predictions made by learners’ intermediate and advanced grammars (Tableau 6.14 and 6.15):

Tableau 6.14 Intermediate grammar (inversion)

<table>
<thead>
<tr>
<th>loser/winner pairs [+focus]</th>
<th>SUBJ</th>
<th>DROP T</th>
<th>PARSE ; AF-RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>K a. [IP Rosa, [VP t, va a estar allí]]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>b. [IP Ø [VP t, va a estar allí] Rosa]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. [IP Ø [VP t, va a estar allí]]</td>
<td>*!</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 6.15 Advanced grammar (inversion)

<table>
<thead>
<tr>
<th>loser/winner pairs [+focus]</th>
<th>DROP T</th>
<th>PARSE ; AF-RT ; SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>K a. [IP Rosa, [VP t, va a estar allí]]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>K b. [IP Ø [VP t, va a estar allí] Rosa]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>c. [IP Ø [VP t, va a estar allí]]</td>
<td>*</td>
<td>*(!)</td>
</tr>
</tbody>
</table>

The L1 English grammar of learners (Tableau 6.14) does not accommodate inverted subject sentences or null subjects, therefore, (a) is the only grammatical choice. Once SUBJECT has been demoted below DROPTOPIC, however, the restructured grammar now accepts either inverted (b) or non-inverted (a) choices. Tableau 6.15c also makes a specific prediction not tested by this study, namely, that focused elements must be parsed; if left unparsed, the resulting candidate would fatally violate more than one constraint on the second tier.

The resulting grammar, therefore, is one in which either SUBJECT or ALIGNFOCUS-RIGHT can (indeed must) be violated, but neither of these violations is now
fatal to the grammar. This means that inverted orders may appear where alignment of focus is occurring, or a non-inverted order may be used (regardless of focus). Either choice is equally acceptable. Therefore, there are two ‘winning’ candidates. The result of this restructuring is a grammar that is ‘relatively free’ in regards to inversion. It is not fully free, because inversion would not be expected to surface in non-focused orders, since there is nothing to motivate the inversion; but, focused orders have the option of violating one constraint or the other, and neither option invokes a higher penalty. Again, the results in this study are consistent with this restructured grammar and with the CDA.

6.4 That-trace

Finally, we consider the observation that English displays an anti-that-trace effect, while Spanish regularly requires that-trace sequences. Some of the results of this study suggest that the grammatical property of that-trace is not salient for learners. Learners, even from the earliest stages, appear willing to accept at least some sentences with que plus a trace; but these learners do not realize the necessity of the use of the complementizer with certain matrix verbs.

Even though learners do not encounter that-trace in their L1, they frequently encounter subjects in subordinate clauses, and the complementizer variably precedes these subjects. The difference between this variability in their L1 and the non-variability of the L2 poses a learning challenge that is distinct but not separate from the acquisition of that-trace. These two distinct pieces of learning are illustrated by 6.4 and 6.5:

(6.4.) a. ¿Quién piensas que ti nos puede ayudar? (Spanish)
b. *Who, do you think that ti can help us? (English)
In 6.4b, the inclusion of *that* in English makes the sentence ungrammatical. In contrast, the Spanish equivalent in 6.4a requires the complementizer, placing the Spanish grammar in direct conflict with English at this point. In 6.5, the inclusion of *that* in English is optional only in English; Spanish requires it. Furthermore, the null subject in 6.5 is not a trace, so we might predict that learners’ grammars may adequately handle this type of item sooner than items such as 6.4 that instantiate *that-trace*.

Several different proposals have attempted to capture cross-linguistic differences in the realization of *that-trace* and complementizer optionality. In Section 2.6 of Chapter 2, two such proposals were discussed — that of Grimshaw (1997), in which the constraints regarding the government of traces (T-Gov and T-Lex-Gov) were employed to account for *that-trace* configurations, and that of Baković (1997), in which complementizer optionality was derived from the interaction of faithfulness and markedness constraints upon differing inputs. Baković’s FAITH[SUB] constraint was based on the idea that some functional features, such as a value of ± for subordination, are a part of the input.

One additional analysis of complementizers is found in Pesetsky (1997), in which cross-linguistic differences are understood, in part, to be the result of the an interaction of two constraints, TELEGRAPH and LEFTEDGE(CP). Pesetsky proposed TELEGRAPH as a syntactic constraint requiring that function words be left unpronounced. As such, it is a constraint that is sensitive to syntactic categories; for example, it is not failed by an overt NP, but it is failed by an overt complementizer. This is in keeping with prevailing
thought regarding economy conditions (Chomsky 1995) or economy of expression (Bresnan 1998), which holds that only elements that are ‘meaningful’ (i.e. contain semantic features) should be expressed.

A different type of economy consideration is at work with the LEFTEDGE(CP) constraint. LEFTEDGE(CP) is an alignment constraint that is discoursal in as much as it requires that information be packaged in a specific way. When a speaker desires to subordinate a thought to a higher matrix thought, arguably a discoursal concern, then a CP is projected to accomplish this information packaging requirement. CPs need not generally be projected for all clauses, as once thought, because the projection of an empty CP adds no additional information and thus violates general principles of economy (Chomsky 1995). When the projection of a CP does provide additional information (e.g. that the utterance following should carry the information-packaging feature of subordination) then a CP has reason to be projected. LEFTEDGE(CP) requires that in such a projection the leftmost overt word be the complementizer, a lexical head instantiating the category and aligning it in a manner similar to other alignment constraints. LEFTEDGE(CP), therefore, is violated when an embedded clause begins with anything other than the complementizer.

Of all the constraints proposed to account for complementizer effects, LEFTEDGE(CP) appears most susceptible to the critique of being ad hoc because it makes reference to a specific projection (CP) and a specific element within that projection (a complementizer). Despite this appearance, LEFTEDGE(CP) seems well motivated both because it effectively accounts for a broad range of data and because variant rankings of LEFTEDGE(CP) with respect to a constraint such as TELEGRAPH yield sets of actual
grammars. As with inversion, three possible rankings of these two constraints are possible, and these rankings each correspond to different human languages. Figure 6.7 shows the possibilities:

Figure 6.7 Ranking TELEGRAPH and LEFTEDGE(CP)

(a) TELEGRAPH » LEFTEDGE(CP) (Chinese)
(b) LEFTEDGE(CP) » TELEGRAPH (French)
(c) TELEGRAPH, LEFTEDGE(CP) (English)

The English grammatical ranking is that of Figure 6.7c, where optionality in the presence of that exists; both I believe John won and I believe that John won are grammatical. If TELEGRAPH outranks LEFTEDGE(CP) (Figure 6.7a), then complementizers are always dispreferred, which is the case with a language such as Chinese. Finally, if LEFTEDGE(CP) outranks TELEGRAPH (Figure 6.7b), then complementizers regularly appear, and the non-complementizer option is dispreferred, as in French.

Pesetsky (1997) also examined variation in the use of complementizers in French and English. He demonstrated that LEFTEDGE(CP) and TELEGRAPH also interact with a recoverability constraint (REC) that requires that the semantic content of unpronounced elements be recoverable from the local context. Pesetsky claimed that this interaction accounted for distribution of the complementizer not only in declarative clauses, but also in simple and complex relative clauses and in embedded questions. For example, Pesetsky posited only a minimal difference between the ranking found in French (RECOVERABILITY » LEFTEDGE(CP) » TELEGRAPH) and the ranking found in English (RECOVERABILITY » LEFTEDGE(CP), TELEGRAPH), namely, that French ranks
LEFTEDGE(CP) over TELEGRAPH, while these two constraints are unranked with respect to each other in English.

This difference in ranking between the grammars of French and English yields a multitude of differences in grammaticality. For example the optionality of LEFTEDGE(CP) and TELEGRAPH in English (Tableau 6.22), but not in French (Tableau 6.23), results in variant outputs with respect to the declarative complementizer. (Examples taken from Pesetsky 1997:157-162; French and English examples are equivalent):

Tableau 6.16  Declarative complementizer (English)

<table>
<thead>
<tr>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tableau 6.17  Declarative complementizer (French)

<table>
<thead>
<tr>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td>!</td>
</tr>
</tbody>
</table>

This variance in ranking also accounts for the difference between French and English in the outcome of simple relative clauses (Tableaux 6.18 and 6.19):

Tableau 6.18  Simple relative clauses (English)

<table>
<thead>
<tr>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

206
Tableau 6.19  Simple relative clauses (French)

<table>
<thead>
<tr>
<th></th>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP qui que je connais]</td>
<td>*</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP qui que je connais]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP qui que je connais]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP qui que je connais]</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

The interaction of constraints here also accounts for the similarity of English and French in the outcome of complex relative clauses (Tableaux 6.20 and 6.21) despite the difference in the constraint rankings of these languages:

Tableau 6.20 Complex relative clauses (English)

<table>
<thead>
<tr>
<th></th>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>the man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP with whom that I danced]</td>
<td>*</td>
<td></td>
<td>*(!)</td>
</tr>
<tr>
<td>K the man [CP with whom that I danced]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>the man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP with whom that I danced]</td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>the man</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP with whom that I danced]</td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau 6.21 Complex relative clauses (French)

<table>
<thead>
<tr>
<th></th>
<th>REC</th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP avec qui que j’ai dansé]</td>
<td>*</td>
<td></td>
<td>*(!)</td>
</tr>
<tr>
<td>K l’homme [CP avec qui que j’ai dansé]</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP avec qui que j’ai dansé]</td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>l’homme</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP avec qui que j’ai dansé]</td>
<td></td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

The broad range of crosslinguistic data that are explained by positing a LEFTEDGE(CP) constraint lends credibility to Pesetsky’s constraint and warrant its consideration for the acquisition of the complementizer properties of Spanish. If Pesetsky’s analysis for French — which yields the result that complementizers regularly appear and the non-
complementizer option is dispreferred — may be extended to Spanish, then we may posit that $\text{LEFTEDGE}(\text{CP})$ is implicated in the necessity of the Spanish complementizer as well.

6.4.1 An initial analysis of that-trace

Pesetsky’s proposal for complementizer optionality initially appears very promising. The main question for this developmental account, however, is whether the differences in the constraint rankings of English and Spanish related to $\text{LE}(\text{CP})$ and $\text{TELEGRAPH}$ find a natural explanation under the proposed learning algorithm (CDA). The analysis here will eventually argue that Pesetsky’s proposal fails in this regard, but the locus of the failure may be important for possible further refinements to the CDA or other learning algorithms, so this section makes an initial attempt at analyzing the acquisitional path related to complementizer optionality and that-trace using Pesetsky’s constraints.

When learners hear questions such as 6.6, where the inclusion of $\text{que}$ is necessary even if it violates that-trace, they begin to realize that a restructuring of their grammar is necessary to account for this new property.

(6.6.) ¿Quién cree que él llamó?
who think-2sg-pres that he called-3sg-pret.
‘Who do you think called?’

We might then assume that, as with the previous demotions, the grammar evaluates winner/loser pairs of candidates against the constraint hierarchy (Tableau 6.22).
The problem with applying the CDA now becomes immediately obvious. The learner grammar needs to demote TELEGRAPH below LE(CP) so that the requirement to omit function words gives way to the requirement for an overt complementizer. But how can CDA do this when both candidates in the initial hierarchy are winners? Without a loser to motivate the activation of demotion, there is no reason why either constraint should be demoted. Demoting TELEGRAPH below LE(CP) does indeed yield the correct results, but the current conception of CDA does not permit such a demotion here.

Nevertheless, if we suspend our judgment regarding the feasibility of such an analysis under CDA, several interesting observations may be made. If we assume that there is some mechanism by which demotion of TELEGRAPH below LE(CP) may take place, then we must note that the mark-data pair in 6.22 involves the operation of an interlanguage stage. For the L1 English to be compared with the Spanish, a fuller tableau undoing previous demotions related to null subjects would be needed. Without these intervening demotions, the non-ranking of TELEGRAPH and LEFTEDGE(CP) would suggest that the complementizer is fully optional in English, even in cases where a that-trace sequence would be generated, and this is not the case. This means that Tableau 6.22 could only represent a stage subsequent to the early stages enabling null topic subjects. In cases where the subject is a non-topic, the Spanish would also have an overt pronoun in place of the trace in Tableau 6.22; this would make the correct prediction for English
— that either of the options in Tableau 6.22 is acceptable if there is an overt subject in
the lower clause.

That dependency of Tableau 6.22 upon earlier reranking regarding null subjects
yields one straightforward explanation of why *that-trace* surfaces later than null subjects
in the learner grammars of the L1 English speakers in this study — *that-trace* sequences
are disallowed in the early grammars because null subjects are disallowed. With the
onset of null subjects, *that-trace* becomes acceptable, and the new learning task is simply
to restrict the optionality involved in the realization of *que*. Tableau 6.23 shows the
constraint demotion that would remove this optionality by moving TELEGRAPH below
LEFTEDGE(CP):

Tableau 6.23  Demotion of TELEGRAPH

<table>
<thead>
<tr>
<th>loser/winner pairs</th>
<th>TEL</th>
<th>LE(CP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winner: ¿Quién crees [CP  que  ti llamó]?</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Winner: ¿Quién crees [CP  ∅  ti llamó]?</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

This demotion would results in the new ranking in Tableau 6.24:

Tableau 6.24  Ranking after the demotion of TELEGRAPH

<table>
<thead>
<tr>
<th></th>
<th>LE(CP)</th>
<th>TEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>¿Quién crees [CP  que  ti llamó]?</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>¿Quién crees [CP  ∅  ti llamó]?</td>
<td>*!</td>
</tr>
</tbody>
</table>

The result of this demotion is that there are now clearly losing and clearly
winning candidates. Violations of LEFTEDGE(CP) are now fatal to the grammar. The
result of this restructuring is that choices between candidates should have a clear winner,
where they previously did not. Such a result finds some support in this study. For
example, in dialogue 9 of the grammaticality judgment task, only 43% of the beginners
choose to include *que*, but this percentage steadily increases, until near-native and native speakers include *que* 100% of the time with the verbs used in this study.\(^{10}\)

As attractive as such an analysis might be, the problem remains that the demotion in Tableau 6.23 is not feasible under the CDA. The application of Pesetsky’s analysis to the acquisition of Spanish in this study has additional limitations. Although his account addresses the issue of complementizer optionality, it does not address the issue of *that-trace*. As was seen above, *that-trace* effects may be derived with the use of Pesetsky’s constraints, but only in concert with other constraints regarding null subjects. This is not a serious limitation, since the necessity of earlier constraint demotions would provide further confirmation of the main argument of this dissertation — that the grammatical properties subsumed under pro-drop are actually acquired at different times and in a certain developmental order — but the inability to use the CDA to account for this demotion means the Pesetsky’s analysis must be discarded in this developmental account.

### 6.4.2 A revised analysis of *that-trace*

Since the account of Pesetsky (1997) encountered difficulties with the operation of the CDA, it may be asked whether the accounts of Grimshaw (1997) or Baković and Keer (2001) fare any better. Grimshaw’s account introduces the constraint requiring that traces be lexically governed (T-LEX-GOV), and Baković and Keer employ this constraint in competition with a faithfulness constraint ensuring that the output for each candidate conforms to that candidates specification for subordination found in the input.

This section argues that Baković and Keer’s analysis, applied to developmental data, operates within the parameters of the CDA. The favored analysis in this dissertation
interprets the developmental data of this study through a two-step learning process, the first step which involves an awareness of the CP status of all Spanish embedded clauses, and the second step which involves standard operation of the CDA.

We begin by assuming that the initial constraint ranking (English) is that found in Figure 6.8a, and the target constraint ranking (Spanish) is that of 6.8b.

Figure 6.8 Rankings yielding that-trace

a. English ranking: T-LEX-GOV » FAITH[SUB]
b. Spanish ranking: FAITH[SUB] » T-LEX-GOV

To move from the ranking in Figure 6.8a to the ranking in Figure 6.8b (Spanish), learners must demote T-LEX-GOV to the stratum below FAITH[SUB], but even prior to this demotion, learners begin to detect a subtle difference between their native English grammars and the grammar of Spanish. The results of this study showed that advanced learners accurately chose the overt complementizer in Spanish for those sentences whose English equivalent demonstrated apparent optionality, but did not accurately handle that-trace sentences until the near-native stage. Learners apparently first realize that sentences such as 6.7c and 6.7d, which appear to be optional in English, require the presence of the complementizer in Spanish (6.7a and 6.7b):

(6.7.) a. Yo creo que vamos a encontrar a alguien. (Spanish)
    b. *Yo creo vamos a encontrar a alguien.
    c. I think we are going to find someone. (English)
    d. I think that we are going to find someone.

The appearance of optionality in the English examples in 6.7, however, may be deceiving. The optionality of 6.7 vanishes if it is assumed that 6.7c and 6.7d are based on differing inputs, one with a [-subordination] feature and one with a [+subordination]
feature, making the structural distinction between 6.7c and 6.7d such that 6.7c is an IP and 6.7d a CP (Doherty 1993, Baković and Keer 2001).

The ranking of T-LEX-GOV » FAITH[sub] in English yields an outcome that reflects the designation of the verbal extended projection in the input: if the verbal extended projection is an IP, no complementizer will appear in the output; if the verbal extended projection is a CP, the complementizer will appear. Since sentences such as those in 6.7 do not contain traces, T-LEX-GOV is vacuously satisfied regardless of the designation given in the input (Tableau 6.25a and 6.25b):

Tableau 6.25a  Complementizer ‘optionality’ in English [+sub]

<table>
<thead>
<tr>
<th>Input: [+sub]</th>
<th>T-LEX-GOV</th>
<th>FAITH[sub]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. I think [IP we are going to find someone.]</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>b. I think [CP that we are going to find someone.]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tableau 6.25b  Complementizer ‘optionality’ in English [-sub]

<table>
<thead>
<tr>
<th>Input: [-sub]</th>
<th>T-LEX-GOV</th>
<th>FAITH[sub]</th>
</tr>
</thead>
<tbody>
<tr>
<td>K a. I think [IP we are going to find someone.]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. I think [CP that we are going to find someone.]</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

In Tableau 6.25a and 6.25b, whether the input calls for an overt complementizer (CP) or no complementizer (IP), the FAITH[sub] constraint eradicates any lack of correspondence between the input and output of the subordinator feature.

Interestingly, transfer of this L1 constraint ranking will result in correct outputs for Spanish input in sentences such as 6.7a (above), but only once learners recognize that the Spanish input normally carries the [+sub] feature. Learners eventually make native-like choices with items such as these, but they first become fully aware of the persistent [+sub] feature at the advanced level. Tableau 6.26 (below) shows that in Spanish, as in...
English, violation of FAITH[SUB] results in ungrammaticality. Once again, T-LEX-GOV is vacuously satisfied.

Tableau 6.26 Obligatory complementizer in Spanish [+sub]

<table>
<thead>
<tr>
<th>[+] sub</th>
<th>T-LEX-GOV</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
</table>
| K  
Yo creo\text{[CP que vamos a encontrar a alguien.]} & & *! |
| Yo creo\text{[IP vamos a encontrar a alguien.]} & & |

Therefore, the initial piece of learning that must take place is recognition that Spanish input displays a consistency not found in English – lower clauses in English may be IPs or CPs, while Spanish lower clauses are consistently CPs. Since this learning relates to an awareness of the input, no constraint reranking is necessary to yield appropriate outputs.\textsuperscript{11}

Even after learners have become aware that Spanish input favors the [+sub] feature, the acquisition of that-trace is not yet complete. Since the L1 English grammar has not yet been required to restructure to handle Spanish inputs, the grammar is still one that does not permit violation of that-trace. Therefore, an English complement clause that has undergone subject extraction will not instantiate the complementizer, regardless of the subordination feature supplied by the input (Tableau 6.24a and 6.24b):

Tableau 6.27a Complementizer absence in English [+sub]

<table>
<thead>
<tr>
<th>input: [+]sub</th>
<th>T-LEX-GOV</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
</table>
| K  
Who do you think \text{[CP ti called]}? & & * |
| Who do you think \text{[CP that ti called]}? & & *! |

Tableau 6.27b Complementizer absence in English [-sub]

<table>
<thead>
<tr>
<th>input: [-sub]</th>
<th>T-LEX-GOV</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
</table>
| K  
Who do you think \text{[CP ti called]}? & & *! |
| Who do you think \text{[CP that ti called]}? & & * |
Since this result is at variance with the target (Spanish) grammar, Constraint Demotion is triggered but, unlike the problem posed by Pesetsky’s initial non-ranking of LE(CP) and TELEGRAPH, here the CDA functions properly because a winner/loser pair may be identified (Tableau 6.28):

Tableau 6.28 Mark-data pair: Advanced (that-trace)

<table>
<thead>
<tr>
<th>loser/winner pairs [+sub]</th>
<th>T-LEX-GOV</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser: ¿Quién, crees [CP t_i llamó]?</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Winner: ¿Quién, crees [CP que t_i llamó]?</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

In Tableau 6.28, the winner mark, T-LEX-GOV, must become dominated by the loser mark, FAITH[SUB]. Therefore, the CDA requires that T-LEX-GOV be demoted to the stratum immediately below that of the loser mark, FAITH[SUB] (Tableau 6.29).

Tableau 6.29 Demotion of T-LEX-GOV

<table>
<thead>
<tr>
<th>loser/winner pairs [+sub]</th>
<th>T-LEX-GOV</th>
<th>FAITH[SUB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loser: ¿Quién, crees [CP t_i llamó]?</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Winner: ¿Quién, crees [CP que t_i llamó]?</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

The resulting ranking is shown in Tableau 6.30:

Tableau 6.30 Ranking after the demotion of T-LEX-GOV

<table>
<thead>
<tr>
<th>[+sub]</th>
<th>FAITH[SUB]</th>
<th>T-LEX-GOV</th>
</tr>
</thead>
<tbody>
<tr>
<td>¿Quién, crees [CP t_i llamó]?</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>K¿Quién, crees [CP que t_i llamó]?</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

As with the first analysis using TELEGRAPH and LE(CP), this demotion now yields losing and winning candidates in keeping with Spanish grammar. Since this ranking converges on the target grammar, no further demotions are needed. This reranking takes learners to the near-native stage. The final grammar is one in which the complementizer
que will be present whenever it is present in the input, even if this results in a violation of T-LEX-GOV. Since lower clauses in Spanish normally include the [+sub] feature, que will regularly appear in all subordinate clauses, as the data from this study also confirms.

6.6 Summary of the analysis

To summarize the findings regarding how the developmental path of learners involves recursive stages of constraint demotion, each stage of the grammar is now presented with a total ranking of all six constraints. This shows that each stage, from beginners through near-native speakers involves new demotions, some of which could not take place until they were preceded by others. The fact, for example, that the demotion resulting in inversion could not precede the demotions resulting in null subjects, has no clear account in a GB-theoretic analysis, but it actually is accounted for under the CDA analysis presented here. Consider the following total rankings and each movement that takes learners to the next stage (Figures 6.9)

Figure 6.9 Constraint demotion and the total ranking

<table>
<thead>
<tr>
<th>Beginners (English): PARSE » SUBJ » DROPT » {AF-Rt, T-LEX-GOV} » FAITH[SUB]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demote PARSE below DROPTOPIC</strong></td>
</tr>
<tr>
<td>Intermediates: SUBJ » DROPT » {PARSE, AF-Rt, T-LEX-GOV} » FAITH[SUB]</td>
</tr>
<tr>
<td><strong>Demote SUBJECT below DROPTOPIC</strong></td>
</tr>
<tr>
<td>Advanced: DROPT » {SUBJ, PARSE, AF-Rt, T-LEX-GOV} » FAITH[SUB]</td>
</tr>
<tr>
<td><strong>Demote T-LEX-GOV below FAITH[SUB]</strong></td>
</tr>
<tr>
<td>Near-native (Spanish): DROPT » {SUBJ, PARSE, AF-Rt} » FAITH[SUB] » T-LEX-GOV</td>
</tr>
</tbody>
</table>
Each stage in Figure 6.9 is represented by one demotion, and each demotion yields a new grammar that is used by learners to progressively move towards the target language. The need to demote PARSE below DROPTOPIC initially confuses beginning learners, and this is reflected in the overgeneralization of subject dropping. Intermediate learners eventually become cognizant of the discourse condition involved with null subjects, and when they demote SUBJECT below DROPTOPIC, the result is not only a refined grammar in regards to null subjects, but also the emergence of grammar that permits inversion of focused constituents. At the advanced levels, learners become aware that Spanish lower clause inputs regularly carry the [+sub] feature. This realization results in the obligatory instantiation of que in sentences where English displays apparent optionality. The final demotion of T-LEX-GOV below FAITH[SUB] permits learners to overcome the grammaticality constraint in their L1 that would normally prohibit that-trace sequences. Each of the demotions in Figure 6.9 is consistent with the CDA, lending support for its application to issues of second language development.

6.7 General conclusions and implications for future research

This dissertation has proposed a new account of the developmental path second language learners take in regards to the grammatical properties traditionally associated with pro-drop. This proposal has been supported by empirical tests of the L2 acquisition of Spanish by native speakers of English. These tests demonstrated that the various elements encapsulated in the term ‘pro-drop’ are acquired by second language learners in a particular developmental order that is predictable from the relative rankings of grammatical constraints in the native and second languages. These results challenge
traditional parameter setting accounts of pro-drop by arguing that the grammatical properties associated with pro-drop are epiphenomena resulting from particular constraint rankings within a grammar, not the switching of a single parameter to a particular, inviolable setting.

The first hypotheses of this dissertation (H1 in Section 4.1 of Chapter 4) held both that the implicational hierarchy of Liceras (1989) would be validated by the results here and also that this hierarchy would be insufficient to distinguish between the initial acceptability of null subjects, inversion and that-trace, and accurate judgments regarding these properties. The implicational hierarchy of Liceras (1989) did not find strong support because that-trace choices surfaced earlier and at a greater rate than expected. Nevertheless, the findings of this study are not inconsistent with Liceras’ hierarchy, and in Section 5.4 of Chapter 5, some reasons were forwarded for maintaining the hierarchy despite its lack of positive evidence here.

The second hypothesis (H2 in Section 4.1 of Chapter 4), that initial acceptability would be distinct from accurate judgments, found strong confirmation. Null subjects surfaced early, but it was not until the late stages of acquisition that these subjects were selected with a native-like frequency and distribution. Inversion presented a bit more complicated case. Learners did not often select inversion; however, choices for inversion occurred more frequently when constituents were focused, and the difference between these conditions is likely due to the demotion of certain syntactic constraints in English that compete with discoursal constraints in Spanish. That-trace choices may have surfaced early, but these choices were not accurately acquired until much later, some time between the advanced and near-native stages.
A third hypothesis (H3 in Section 4.1 of Chapter 4) held that some L2 learners would eventually (in the later stages of acquisition) converge on native-like usage of null subjects, inversion and that-trace, and that this convergence would come as a result of a sensitivity to the discoursal constraints of the language. This hypothesis dealing with ultimate attainment was largely confirmed (although the evidence for convergence in regards to inversion is less clear). As the analysis in this chapter demonstrated, final convergence in two cases was the result of sensitivity to discoursal constraints: for null subjects, convergence came through an awareness of the proper ranking of DROPTOPIC with respect to syntactic constraints and for inversion, convergence came as a by-product of the reranking of SUBJECT of DROPTOPIC that created a new non-ranked relationship between SUBJECT and ALIGNFOCUS-RIGHT. Each case involved demotion of syntactic constraints in relation to discoursal constraints in the target language.

The results here argue against accounts in which autonomous applications of syntax fail to admit interactions between syntax and discourse. The developmental path taken by L2 learners is best characterized in terms of the interaction between these two components of grammar, and the precise path that learners take confirms the final hypothesis — that these results have a natural interpretation under the operation of the Constraint Demotion Algorithm of Tesar and Smolensky (2000). There is strong evidence that the interaction between syntax and discourse for L2 learners is not parameterized — lower-ranked constraints remain operative within the language and higher-ranked constraints are sometimes violated. Both of these circumstances challenge those formulations of pro-drop that are built on inviolable principles.
One contribution to linguistic theory that this study makes is a clearer statement of the facts related to the path L2 learners take as they acquire the features of a ‘pro-drop’ language. Perhaps this contribution may advance linguistic theory not only by specifically providing empirical support for the Constraint Demotion Algorithm (CDA) proposed by Tesar and Smolensky (2000), but also by reinforcing the soundness of an Optimality Theoretic approach to issues in language learning more generally.

The application of OT to second language learning represents a new and potentially productive line of inquiry. The results here provide support for viewing the variant outputs of interlanguage grammars as differences in the rankings of syntactic and discoursal constraints. The results also leave many questions unanswered, however, and suggest a future research program devoted to exploring how constraint interactions between various levels of grammatical knowledge are used by learners in the acquisition of a second language. There is a clear need for future research in this area, perhaps beginning with further improvements on the research design used here, but also extending to the study of other language groups and other grammatical properties. My hope is that this project has provided some new insights into interfaces between syntax and discourse in the acquisition of Spanish learner grammars, and that these insights will lead to further investigations that will advance our understanding of second language acquisition.
Notes

1 Some researchers (e.g. Broekhuis and Dekkers 2000) prefer the ‘tied’ or equal ranking option, accepting A»B and B»A, but the option of nonranking with respect to each other is used by Tesar and Smolensky (2000) and is followed here. One difficulty with accepting both rankings, aside from the challenge it would seem to pose to OT conceptually, is that such the learner would continue to receive evidence contrary to each of the rankings, creating an instability in which the grammar demotes one constraint (A) below the other (B) and then demotes that constraint (B) back below (A) in an endless loop. While this is theoretically possible, the failure for the CDA to converge in such instances would be a less-than-welcome result.

2 The possibility of constraints with equal ranking is not without certain theoretical difficulties. For example, some of the formal results demonstrating constraint demotion are based upon totally ranked hierarchies, and likewise, most conceptions of OT regard target grammars to be totally ranked hierarchies. Tesar and Smolensky (2000:48) concede that stratified hierarchies present the possibility that constraint demotion will fail to converge, rearranging constraints with each piece of conflicting data; nevertheless, it is possible to distinguish, as Tesar and Smolensky do, between the presence of stratified hierarchies in learner grammars and their presence in target grammars. The learning algorithm may operate in the larger space (i.e. that of stratified hierarchies) while target grammars may be totally ranked hierarchies, or on the verge of achieving this total ranking. One way of arriving at this result is to make the endpoint of learning one in which hierarchies are not fully ranked, because the learning algorithm is left without sufficient evidence to further refine the grammar into a fully ranked hierarchy (Tesar and Smolensky 2000:49). Another way constraint demotion may learn stratified hierarchies is to assume ‘multiple optimal outputs’ (Tesar and Smolensky 2000:50) that do not involve an identical set of marks. Nevertheless, as Tesar and Smolensky (2000:50) state, ‘achieving ties for optimality between forms that incur different marks is always a delicate matter’. Therefore, although the analysis used in this chapter assumes that stratified hierarchies exist in the native, learner, and target grammars, it would be useful to pursue further research either into constraint interactions that might derive the same results without the use of nonranked or tied constraints, or into further refinements of the learning algorithm that would more efficiently handle stratified hierarchies.

3 Since PARSE does not dominate SUBJECT in Spanish, violations of PARSE may appear in a broader range of contexts than English. The most common example is that PARSE no longer prohibits null topic-connected subjects, because it is dominated by DROPTOPIC. Unlike English, in Spanish, PARSE, SUBJECT, or both constraints may sometimes be violated without incurring a fatal violation. In Ecuadorian Spanish, for example, topic-connected objects may drop (cf. Suñer and Yépez 1988), violating PARSE, but not SUBJECT:

A: ¿Comió Juan el pan?  A': Did John eat the bread?
B: Sí, Juan comió Ø.  B': *Yes, John ate Ø.
4 The analysis that follows is not strictly dependent upon the validity of the assumption that PARSE and SUBJECT are unranked in Spanish. If PARSE does dominate SUBJECT also for Spanish, then the only required modification to the current analysis would be to first demote PARSE to a stratum below DROP_TOPIC, but above ALIGN_FOCUS-RIGHT, and then demote SUBJECT below PARSE to the stratum of ALIGN_FOCUS-RIGHT. This possibility does yield the same results as the current analysis.

5 The inclusion of ALIGN_FOCUS-RIGHT in this tableau is to show the stratum below DROP_TOPIC which may become the landing site for demotions of constraints initially dominating DROP_TOPIC.

6 It would be useful to investigate how naturalistic language acquisition compares to instructed learning in this regard. The possibility that learners are simply reflecting some instructional effect would find greater support if learners in an uninstructed, naturalistic acquisition environment do not overgeneralize null subjects in the same manner as the learners in this study did.

7 Alignment constraints could result in leftward, as well as rightward, movement of constituents. The constraint discussed here is specified as ALIGN_FOCUS-RIGHT. Movement to this position is not the only possible mechanism for focusing elements. For example, focus in both English and Spanish may occur in situ by placing prominent intonational stress on the focused constituent. Also, fronting constituents to indicate contrastive focus, as in Carrots, I like provides another focusing strategy. These processes are related to the focus discussed in this chapter, in as much as they likely involve an interaction between differing levels of grammar (phonology, discourse, or syntax), but they are distinct from the focus strategy discussed here, and the constraints that account for these variant strategies are a topic for further research.


9 Pesetsky (1997:159) also shows that interactions between LEFT_EDGE(CP) and TELEGRAPH account for the absence of pronounced complementizers in verb-second orders in German.

10 There are verbs in Spanish that do not require the use of que for subordination, but their number is very limited. The account provided here does not explain why these verbs have special status, but it would explain how verbs that provide the [-sub] feature to the input would be realized in the output without a complementizer. The FAITH[SUB] constraint must require the output to conform to the lexical features of the verb (subordination being one of these features). This places the locus of variation in the requirement for the complementizer in the lexicon. The grammar interacts with the input supplied by the lexicon by either attempting to maintain the input (via faithfulness constraints) or by attempting to make the input conform to some syntactic principle (via markedness constraints)
It should be noted that some researchers have reservations about OT analyses that derive systematic differences between languages through the use of differing inputs. The analysis here, as with the analysis of Baković and Keer (2001), may be perceived to be at variance with the standard OT principle of the ‘Richness of the base’ (Prince and Smolensky 1993, section 9.3). This principle states that ‘the set of possible inputs to the grammars of all languages is the same’. Tesar and Smolensky (2001:30) assert that ‘all systematic properties of the lexicon arise indirectly from the grammar, that delimits the inventory from which the lexicon is drawn’. However, even if we maintain that crosslinguistic variation is primarily a result of the operation of a grammar, the nature of the input a grammar uses in its operation is still a matter of lively debate, particularly in OT syntax. While most agree that predicate/argument structure is a part of the input, whether or not functional features may be included has not yet been resolved. The analysis here, following Baković and Keer (2001), assumes that these functional features are indeed available to the syntactic learner and used in the operation of the grammar.
References


Angluin, Dana. 1978. Inductive inference of formal languages from positive data. Information and Control 45.117-35.


Fries, Charles. 1948. As we see it. Language Learning 1.12-16.


Lakshmanan, Usha. 1986. The role of parametric variation in adult second language acquisition: A study of the 'pro-drop' parameter. PALM 2.55-64.


