Romance Conjugational Classes:
Learning from the Peripheries

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By
Angelo Roth Costanzo, M.A.
Graduate Program in Linguistics

The Ohio State University
2011

Dissertation Committee:
Brian D. Joseph, Advisor
Dieter Wanner
Daniel Collins
Copyright by
Angelo Roth Costanzo
2011
ABSTRACT

There is a difference between what linguists do and what speakers do. The goal of the linguist is to describe a language and to come up with some sort of system (be it syntactic, morphological, phonological, etc.) that accounts for the phenomena observed in the language. On the other hand, speakers are not concerned with any of this and their goal is to understand and be understood.

The generalizations that linguists make about language are often times too broad and the desire to reach an “elegant” solution sometimes comes at the cost of the ability to account for all the data (as well as the psychological reality of their solution). Often problematic aspects of the language are left aside and deemed “irregular”, somehow being “outside” of the system being accounted for. Moreover, the data most linguists use describing linguistic phenomena (especially with well-studied languages) come from a standardized form of the language, which can be at times artificial and divorced from the way that people actually speak. A close, detailed examination of data shows that the generalizations that speakers make are not necessarily broad, and are usually local in nature (Joseph & Janda 1986, Joseph 1997).

I argue that the local nature of linguistic generalizations can be seen in the Latin and Romance conjugational class systems, the traditional views of which often try to boil down the differences to somewhere between three and five large conjugational classes. Given that there are many more than between three and five surface patterns of verb conjugation for Latin or any Romance language, it is necessary that some of these classes need to be broken down into several subclasses in order to account for the fact
that different verbs normally considered to be of the same conjugational class often show inflectional patterns that differ in matters of detail. While the establishment of “subclasses” implies that there is a smaller generalization at play (verbs of subclass A are different from verbs of subclass B, even if they belong to the same superordinate class), I argue that the whole notion of “subclass” is unnecessary, and that it is just an artifact of attempts to make an elegant solution.

The data presented in this dissertation are from Romance varieties often seen as “peripheral” and come from dialect atlases, the internet, traditional grammars, and fieldwork undertaken in Sicily and Macedonia. These data show that speakers of Romance languages (as well as speakers of Latin) conjugate verbs that follow a large number of patterns. While it may not clear how these groups are organized mentally, a situation where different patterns that are clustered around features they share provides a more realistic situation than one of a small number of classes that have extensive subbranching.
DEDICATION

For Montedoro.
ACKNOWLEDGEMENTS

First, I would like to thank my dissertation committee. To Brian Joseph, for his patience and encouragement throughout this process, for convincing me that I could finish this dissertation, and for providing a great deal of support during the last six years. To Dieter Wanner, for always being willing to meet with me, for asking challenging questions, and for helping me through some rough parts of this dissertation. To Dan Collins, for providing incredibly helpful comments and for answering my frequent questions about Slavic morphology.

To Hope Dawson, for serving on several of my committees over the years. To Don Winford and Beth Hume, for fantastic courses and support throughout my years at OSU. To Peter Trudgill, for helping me remember that I really like linguistics. To Eric Baković, Sharon Rose, and Amalia Arvaniti, for helping me start my linguistics career on the right track. To Maria–Rosa Lloret, for letting me take her Catalan phonology course at the Universitat de Barcelona and for providing invaluable help with this dissertation. To Judith Tonhauser, for help in planning my fieldwork. Things definitely would not have gone as smoothly without her advice. To Victor Friedman, for helping me prepare for my trip to Macedonia. To Eric Hamp, for inspiring me to do fieldwork in the Balkans. To Martin Maiden, Steve Fondow, Marc–Olivier Hinzelin, James Augerot, Andrea Sims, Catalin Anghelina, and Andrei Crețu, for helpful comments and suggestions.

I am thankful for everyone in Montedoro and Bitola for making me feel welcome and helping me in the fieldwork process. To Genoveffa Chiarelli and Raffaele Costanzo, for giving me a place to stay and for treating me incredibly well. To Riccardo Montagna,
for helping me adjust to life in Montedoro. To Lillo Paruzzo, for taking time out of his
schedule to meet with me almost daily and for many stories about Montedoro. To
Marjan Markovik’ for meeting with me in Skopje. To Jovan Naumovski, Marija Paligora,
Nico Popnicola, and Tasha Jovanovski and family, for providing invaluable help while I
was in Bitola. To the Comună a armănjlor “frats manakia” in Bitola, for welcoming me to
a meeting of their organization. To Dina Cuvata, Goran Kostov, and the Unia ti cultură–
a armănjlor dit machidunii, for meeting with me and providing me with a wealth of
Aromanian material. To the people of Bitola, Malovište, and especially Gopeš, for being
so amazingly hospitable.

I also thank my fellow linguists for support and friendship over the years. To
David Durian, for being an awesome officemate and neighbor. To Julia Papke, for
helping me adjust to life in Columbus and for many delicious meals. To Ila Nagar, for
helping make Oxley 218 a hilarious place. To Nadia El–Yousseph, Grant McGuire, and
Katie Woznicki, for getting me out of my apartment and for hanging out at the Dube. To
Kevin Gabbard and Josh Pennington, for hilarious Kosovo adventures.

To my parents, for unconditional support. To Jo and Tony Galante, for inspiring
me to go to Montedoro. To my brother and sister, for always being there for me. To my
in–laws, for making me feel at home in Columbus. Most of all, I thank my wife, Julia, for
being my best friend, for listening to me complain without getting too frustrated, for
putting up with me talking constantly about Romance verb morphology for the past
several years, and for believing that I could finish this dissertation when I did not believe
in myself.
VITA

2004.......................... B.A. Linguistics, University of California, San Diego
2008.......................... M.A. Linguistics, The Ohio State University
2004–present............... Graduate Teaching Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Linguistics
Minor Fields: Romance Linguistics, Balkan Linguistics
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................. i
DEDICATION .................................................................................................................. iii
ACKNOWLEDGEMENTS ............................................................................................... iv
VITA ............................................................................................................................... vi
FIELDS OF STUDY ........................................................................................................ vi
TABLE OF CONTENTS ............................................................................................... vii
LIST OF TABLES ........................................................................................................... xii
LIST OF FIGURES ......................................................................................................... xvii
SIGNS AND ABBREVIATIONS : GRAMMATICAL ......................................................... xvii
SIGNS AND ABBREVIATIONS : LANGUAGES ............................................................... xviii

## CHAPTER 1 : PRELIMINARIES ............................................................................... 1

1.0. Introduction .......................................................................................................... 1

1.1. Generalizations of Speakers, Generalizations of Linguists .................................... 3

1.1.1. Generalizations ................................................................................................. 3

1.1.2. Paradigm “leveling” .......................................................................................... 5

1.1.3. Generalizations across paradigms ...................................................................... 8

1.1.4. Extension of 1SG.PRES.IND.ACT –m in West & South Slavic .............................. 10

1.2. On Inflectional Classes .......................................................................................... 14

1.2.1. Morphology, morphologization, etc. ................................................................. 14

1.2.2. Inflectional classes ............................................................................................ 18

1.2.3. Structure of Inflectional Classes/Patters ........................................................... 20

1.2.3.1. Dressler’s framework .................................................................................... 20

1.2.3.2. Issues with Dressler’s framework ................................................................. 22

1.3. A Different Take on Conjugational Patterns ........................................................ 24

1.3.1. Defining different conjugational patterns ......................................................... 24

1.3.1.1. All details matter ......................................................................................... 25

1.3.1.2. Distinguishing “patterns” ............................................................................. 26

1.3.1.3. Phonological predictability ......................................................................... 27

1.3.1.4. Branching and subclasses .......................................................................... 30

1.3.1.5. Ignoring history ......................................................................................... 32

1.3.2. Organization of patterns ................................................................................. 34
CHAPTER 6 : ON SOME ISSUES IN SICILIAN AND AROMANIAN VERB CONJUGATION

6.0. Introduction ........................................................................................................... 172
6.1. Methods ............................................................................................................... 172
   6.1.1. Fieldwork ...................................................................................................... 172
   6.1.2. Speaker Recruitment .................................................................................... 173
   6.1.3. Limitations .................................................................................................... 173
6.2. Background on Montedoro, Sicilian Dialects, and Sicilian verb conjugation........ 172
LIST OF TABLES

Table 1.1: Evolution of PRES paradigm of Lat. POSSEUM in Sp./DR ....................................................... 6
Table 1.2: Evolution of PRES.IND paradigm of Lat. POSSEUM in It./Ptg. .................................................... 6
Table 1.3: PIE/Latin PRES.IND paradigm of PIE *wel-/ Lat. volō ............................................................ 7
Table 1.4: Four verbs with unexpected 1SG – y (Wanner 2006) ................................................................. 8
Table 1.5: Athematic/Thematic OCS 1SG.PRES. ....................................................................................... 10
Table 1.6: OCS Athematic/Thematic in 1SG/2SG ..................................................................................... 12
Table 1.7: 2SG.PRES. forms of ‘give’ and ‘eat’ in S. & W. Slavic ............................................................. 12
Table 1.8: 2SG.PRES. forms of ‘be’ in S. & W. Slavic ............................................................................. 13
Table 1.9: Forms of Lat. PORT ‘carry’, LAUDO ‘bring’ ........................................................................... 27
Table 1.10: Outcomes of PIE diaspirate root in Skt, Grk, OE ................................................................. 28
Table 1.11: Spanish stem-changing verbs ............................................................................................... 31
Table 1.12: Forms of Skt. √budh– ‘know’ .............................................................................................. 33
Table 1.13: Sanskrit present classes (traditional view) ........................................................................... 40
Table 1.14: Aorist stem classes (traditional view) .................................................................................. 45
Table 1.15: Present class/aorist class correspondences ........................................................................ 46
Table 2.1: Latin verb categorization .................................................................................................. 52
Table 2.2: Common (oversimplified) picture of Latin verb conjugation ................................................. 53
Table 2.3: Principal parts of Lat. AMO ‘love’ ......................................................................................... 54
Table 2.4: Traditional 3rd vs. CAPIO class vs. traditional 4th ............................................................ 58
Table 2.5: Perfect indicative across traditional (present) conjugational class boundaries .......... 61
Table 2.6: Traditional Latin 1st conjugation verbs, different PERF patterns ........................................ 62
Table 2.7: Alternative Latin perfect types ............................................................................................ 64
Table 2.8: Non 3rd conjugation with alternative perfect stem formation .............................................. 64
Table 2.9: Generalizations of correspondences between perfect and participial stems ....................... 65
Table 2.10: Hale & Buck’s (1903) breakdown of the traditional Latin 3rd conjugation ................. 67
Table 2.11: Nasal infix distribution (shaded cells indicate presence of the nasal) .............................. 68
Table 2.12: Lat. verbs in – no (pres. stem from PIE *CV–n–H) .............................................................. 69
Table 2.13: Present/Perfect of inchoative verbs ...................................................................................... 70
Table 2.14: Defective verb (no morphological present tense) ............................................................... 71
Table 2.15: Latin deponent verbs vs. PASS.PRES conjugation .............................................................. 72
Table 3.1: Leveling of present paradigm of Lat. POSSE in Sp. & DR ....................................................... 77
Table 3.2: Deponent verbs → predominant pattern .............................................................................. 79
Table 6.4: PRES.IND paradigm of Sic. \textit{fumári} 'smoke' ............................................................. 180
Table 6.5: PRES.IND paradigms of Sic. \textit{jittárë} 'toss' and \textit{tirárë} 'throw' .................................................. 180
Table 6.6: PRES.IND paradigm of Sic. \textit{parirë} 'seem' ................................................................. 180
Table 6.7: PRES.IND paradigms of Sic. \textit{cúrrirë} 'run' and \textit{mintirë} 'put' .............................................. 181
Table 6.8: PRES.IND paradigms of Sic. \textit{léggirë} 'read' and \textit{durmírë} 'sleep' .............................. 181
Table 6.9: Montedorese \textit{ridir} 'laugh' ......................................................................................... 183
Table 6.10: Montedorese \textit{diri} 'laugh' ......................................................................................... 183
Table 6.11: PRES.IND paradigm of Sic. \textit{capirë} 'understand' ...................................................... 184
Table 6.12: PRES.IND paradigm of Sic. \textit{patirë} 'suffer', speakers 1,2 ................................................. 185
Table 6.13: PRES.IND paradigms of Sic. \textit{jírë} 'go' and \textit{éssirë} 'be' .................................................... 185
Table 6.14: 1SG/3PL.PRES.IND of “irregular” Sicilian verbs ............................................................... 186
Table 6.15: PERF.IND paradigms of Sic. \textit{cantárë} 'sing' and \textit{durmirë} 'sleep' ............................. 188
Table 6.16: Typical Sic. dual–stem perfect pattern .............................................................................. 188
Table 6.17: PERF.IND paradigm of Sic. \textit{muvirë} 'stay' ......................................................................... 189
Table 6.18: PERF.IND paradigms of Sic. \textit{dari} 'give' and \textit{stari} 'be' ...................................................... 189
Table 6.19: PERF.IND paradigm of Sic. \textit{scrivirë} 'write', speaker 1 ................................................. 190
Table 6.20: PERF.IND paradigm of Sic. \textit{scrivirë} 'write', speaker 2 ................................................... 190
Table 6.21: PERF.IND paradigm of Sic. \textit{putirë} 'be able', speaker 1 .................................................... 191
Table 6.22: PERF.IND paradigm of Sic. \textit{putirë} 'be able', speaker 2 .................................................... 191
Table 6.23: Examples of unpredictable stem linkages ...................................................................... 192
Table 6.24: Examples of lanachiesc–Vlahu’s 4–conjugation system in AR ............................................. 198
Table 6.25: PRES.IND paradigm(s) of AR \textit{duc} 'take' (lanachiesc–Vlahu 2001) .............................. 200
Table 6.26: PRES.IND paradigm of AR \textit{mác} 'eat' ................................................................................. 201
Table 6.27: PRES.IND paradigm of AR \textit{acumprü} 'buy' .................................................................. 202
Table 6.28: PRES.IND paradigm of AR \textit{cântu} 'sing' ............................................................................ 202
Table 6.29: PRES.IND paradigm of AR \textit{intrü} 'enter' ............................................................................. 202
Table 6.30: PRES.IND paradigms of AR \textit{scol} 'beg' and \textit{aplec} 'get up' ............................................. 203
Table 6.31: PRES.IND paradigm of AR (â)ntreb 'ask'; speakers 1,4 .................................................... 203
Table 6.32: PRES.IND paradigm of AR \textit{portu} 'carry'; speakers 1,3 .................................................... 204
Table 6.33: PRES.IND paradigms of AR \textit{gioc} 'dance' and \textit{scol} 'lift'; speaker 1 .............................. 204
Table 6.34: PRES.IND paradigm of AR \textit{vom} 'vomit', speaker 4 .......................................................... 205
Table 6.35: PRES.IND paradigms of AR \textit{ascundu} 'hide' and \textit{dormu} 'sleep' ........................................ 205
Table 6.36: PRES.IND paradigms of AR \textit{dau} 'give' and \textit{vôi} 'want' ...................................................... 206
Table 6.37: PRES.IND paradigm of AR \textit{crescu} 'grow' ....................................................................... 206
Table 6.38: PRES.IND paradigm of AR \textit{pistisepscu} 'believe' ............................................................. 207
Table 6.39: PRES.IND paradigm of AR \textit{cos} 'sew' .......................................................................... 207
Table 6.40: PRES.IND paradigm of AR \textit{armân} 'remain' ................................................................. 208
Table 6.41: PRES.IND paradigm of AR \textit{zbrâscu} 'speak' ................................................................. 208
Table 6.42: PRES.IND paradigm of AR \textit{zbrescu} 'speak' ................................................................. 209
Table 6.43: PRES.IND active paradigm of AR *lucredz* ‘work’ ...................................................... 209
Table 6.44: PRES.IND paradigm of AR *construedz* ‘build’ ............................................................. 209
LIST OF FIGURES

Figure 1.1: *differ* and *different*, earlier English ......................................................... 35
Figure 1.2: *differ*, *different*, comparatives, Modern English ........................................ 35
Figure 2.1: Dressler’s (2003) view of the Latin 1st conjugation ........................................ 56
Figure 4.1: DR conjugations (with branching) ..................................................................... 100
Figure 4.2: Adaptation of *Atlasul lingvistic român pe regiuni, muntenia și dobrogea* (1996),
           harta 100, charting Std. DR *aiurește* ‘rant–3SG.PRES.IND.ACT’ .......................... 115
Figure 5.1: A traditional view of Catalan conjugational classes ........................................ 138
Figure 5.2: Std. *cuso*, adapted from ALC, Map 550 ..................................................... 151
Figure 5.3: Std. *dormo*, adapted from ALC, Map 665 .................................................. 152
Figure 5.4: Extent of variants [kántuk] and [dɔ́ɾmuk] in Northeastern Catalunya (Girona). .... 158
           Partially adapted from ALC maps 328, 665 ......................................................... 158
Figure 7.1: Traditional view of DR 1st & 4th conjugations .............................................. 217
Figure 7.2: Alternative view, DR *a*-stems & *i*-stems .................................................. 218
Figure 7.3: Traditional view, It. *tenere*, *venire*, & verbs in *-ere/-ire* ............................ 219
Figure 7.4: Alternative view, It. *tenere*, *venire*, & verbs in *-ere/-ire* ............................ 220
<table>
<thead>
<tr>
<th>Sign</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1st person</td>
<td>IND indicative</td>
</tr>
<tr>
<td>2</td>
<td>2nd person</td>
<td>INF infinitive</td>
</tr>
<tr>
<td>3</td>
<td>3rd person</td>
<td>MASC masculine</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
<td>NEUT neuter</td>
</tr>
<tr>
<td>ACT</td>
<td>active</td>
<td>NOM nominative</td>
</tr>
<tr>
<td>AOR</td>
<td>aorist</td>
<td>PASS passive</td>
</tr>
<tr>
<td>COND</td>
<td>conditional</td>
<td>PART participle</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
<td>PERF perfect</td>
</tr>
<tr>
<td>DU</td>
<td>dual</td>
<td>PL plural</td>
</tr>
<tr>
<td>FEM</td>
<td>feminine</td>
<td>PPP past passive participle</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
<td>PRES present</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
<td>PRFV perfective</td>
</tr>
<tr>
<td>GER</td>
<td>gerund</td>
<td>REF reflexive</td>
</tr>
<tr>
<td>IMP</td>
<td>imperfect</td>
<td>SG singular</td>
</tr>
<tr>
<td>IMPV</td>
<td>imperfective</td>
<td>SUBJ subjunctive</td>
</tr>
</tbody>
</table>
## SIGNS AND ABBREVIATIONS: LANGUAGES

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Language</th>
<th>Abbreviation</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGrk</td>
<td>Ancient Greek</td>
<td>Occ</td>
<td>Occitan</td>
</tr>
<tr>
<td>Alb</td>
<td>Albanian</td>
<td>OCS</td>
<td>Old Church Slavonic</td>
</tr>
<tr>
<td>AR</td>
<td>Aromanian</td>
<td>PIE</td>
<td>Proto–Indo–European</td>
</tr>
<tr>
<td>BCS</td>
<td>Bosnian–Croatian–Serbian</td>
<td>Pol</td>
<td>Polish</td>
</tr>
<tr>
<td>Bg</td>
<td>Bulgarian</td>
<td>Prov</td>
<td>Provençal</td>
</tr>
<tr>
<td>Blr</td>
<td>Belorussian</td>
<td>Ptg</td>
<td>Portuguese</td>
</tr>
<tr>
<td>Cat</td>
<td>Catalan</td>
<td>Rms</td>
<td>Romansch</td>
</tr>
<tr>
<td>Cz</td>
<td>Czech</td>
<td>Rus</td>
<td>Russian</td>
</tr>
<tr>
<td>DR</td>
<td>Daco–Romanian</td>
<td>Sard</td>
<td>Sardinian</td>
</tr>
<tr>
<td>Eng</td>
<td>English</td>
<td>Sic</td>
<td>Sicilian</td>
</tr>
<tr>
<td>Fr</td>
<td>French</td>
<td>Skt</td>
<td>Sanskrit</td>
</tr>
<tr>
<td>Frl</td>
<td>Friulian</td>
<td>Sp</td>
<td>Spanish</td>
</tr>
<tr>
<td>Hng</td>
<td>Hungarian</td>
<td>Svk</td>
<td>Slovak</td>
</tr>
<tr>
<td>IR</td>
<td>Istro–Romanian</td>
<td>Svn</td>
<td>Slovene</td>
</tr>
<tr>
<td>It</td>
<td>Italian</td>
<td>Trk</td>
<td>Turkish</td>
</tr>
<tr>
<td>Lat</td>
<td>Latin</td>
<td>Ukr</td>
<td>Ukrainian</td>
</tr>
<tr>
<td>MGrk</td>
<td>Modern Greek</td>
<td>Ved</td>
<td>Vedic</td>
</tr>
<tr>
<td>Mk</td>
<td>Macedonian</td>
<td>Veg</td>
<td>Vegliot</td>
</tr>
<tr>
<td>MR</td>
<td>Megleno–Romanian</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 1:
PRELIMINARIES

1.0. INTRODUCTION

There is a difference between what linguists do and what speakers do. The goal of the linguist is to describe a language and to come up with some sort of system (be it syntactic, morphological, phonological, etc.) that accounts for the observed phenomena. On the other hand, speakers are not concerned with any of this, and their main concern is to understand and be understood. As Andersen (1973) puts it, “it is worth noting that the language learner's goal is the formation not of a specific (‘true' or 'optimal') grammar, but only of a grammar which in some way conforms to the observed data”.

The generalizations that linguists make about language are often times too broad and the desire to reach an “elegant” solution sometimes comes at the cost of the ability to account for all the data, as well as the psychological reality of their solution. Often problematic aspects of the language are left aside and deemed “irregular”, somehow being “outside” of the system being accounted for. Moreover, the data most linguists use in describing linguistic phenomena (especially with well-studied languages) come from a standardized form of the language, which can be at times artificial and divorced from the way that people actually speak. A close, detailed examination of data shows that the generalizations that speakers make are not necessarily broad or widespread; rather they are more commonly local in nature (Joseph & Janda 1986, Joseph 1997).
I argue that the local nature of linguistic generalizations can be seen in the Latin and Romance conjugational class systems, the traditional views of which often try to boil down the differences in verb conjugation to somewhere between three and five large conjugational classes. Given that there are many more than between three and five surface patterns of verb conjugation for Latin or any Romance language, it is necessary that some of these conjugational classes need to be broken down into several “subclasses” in order to account for the fact that different verbs normally considered to be of the same conjugational class often show inflectional patterns that differ in matters of detail. While the establishment of subclass implies that there is a smaller generalization at play here (verbs of subclass A are different from verbs of subclass B, even if they belong to the same superordinate class), I argue that the whole notion of “subclass” unnecessary, and it is just an artifact of attempts to make an elegant solution. However, these so-called “elegant” solutions are often either guilty of leaving data out or of having surface “elegance”, with fragmentation right below the surface.

The data presented in this dissertation show that speakers of Romance languages (as well as speakers of Latin) conjugate verbs that follow a large number of patterns. And while it may not clear how these groups are organized mentally, a situation where different patterns that are clustered around features they share provides a more realistic situation than one of a small number of classes that have extensive subbranching. Moreover, this “clustering” can be different for different speakers of the language (or even vary according to style), given that individual speakers often times show quite individual behavior in verb conjugation when compared to other speakers of the same variety.

Following Joseph (1997b), this dissertation aims to move away from the “core” and instead, focuses on examining the “periphery”. In the “core”, generalizations appear to be more broad. However, in the “periphery”, these generalizations start to break
down, and these marginal phenomena give us a better insight into how speakers generalize and how generalizations work. My preference for focusing on the peripheral goes beyond this; I have also chose to focus on Romance languages that are considered peripheral as they are not given the same attention as the more frequently studied members of the family.

1.1. GENERALIZATIONS OF SPEAKERS, GENERALIZATIONS OF LINGUISTS

1.1.1. Generalizations

This study examines the role and extent of generalizations that speakers in the language use, with particular attention to inflectional verb morphology. The very notion of “generalization” implies a broad scope over a wide range of data, but here I argue that the scope of generalizations is actually much more local in nature, which can be seen when all of the details in the data are sorted out. Normally, when a generalization seems widespread, it is the result of successive “local generalizations” that give the impression of a widespread generalization after the fact.

The generalizations that linguists often make about the behavior of speakers are often too broad in scope and often gloss over small details in the language. For example, as discussed later in this chapter, the common view of the Sanskrit verb system is that there are ten present classes (i.e., ten different ways in which the present stems are formed from roots via a number of morphological processes). However, there are small details (or “exceptions”) within each of the classes. Sometimes linguists leave these details out, but speakers cannot leave this information out. So, for linguists, perhaps the solution that makes most sense is to propose ten verb classes for Sanskrit. However, there is no indication that this is relevant for the speaker, as they exhibit many more patterns than the commonly described ten (see the end of this chapter for more on Sanskrit verb categorization).
Besides ignoring the synchronic facts, often times when linguists claim that a generalization is widespread, its labeling as such is misleading and can be seen in close examination of the facts. The local nature of generalizations can be seen in analogical change (Joseph & Janda 1988, Joseph 2010). For example, the numbers ‘seven’ and ‘eight’ are normally reconstructed in Proto–Indo–European as *septṃ (cf. Lat. SEPTEM, Skt. sapta) and *ökṭā (cf. Lat. OCTO, Skt. aṣṭā), respectively. The forms found in Ancient Greek are expected from these starting points: hepta ‘seven’ and oktā ‘eight’. However, in the Elean dialect of Greek, the resulting outcome of PIE *ökṭā is hoptā. This clearly is not the result of regular sound change, as PIE *k is never expected to give Greek p and Greek ḥ is typically from PIE initial *s-, as seen in the word for ‘seven’, which is not found in the reconstructed word for ‘eight’ (i.e., there is no reason to reconstruct *sökṭā other than this one Greek dialect form). The most likely answer to the origin of this variant form is that in this dialect, the Greek number ‘eight’ was influenced by the number ‘seven’. In allowing ‘eight’ to influence ‘seven’, Elean Greek speakers made a link between the two, in essence connecting them and thus making some sort of generalization over these two forms. Clearly, this is not be considered a wide-spread generalization, as it only affects these two forms (there is not anything saying that “all digits above six accept this consonantism”). This generalization is particular and detailed, and can be described as “local” in scope.

Cases similar to that described above, I argue, are not exceptional. Speakers do generalize ‘locally’, and in some cases the generalization remains local (such as in the example above, where there was a local generalization over two forms that did not spread further whatsoever). Definitely, there are cases where a widespread

---

1 This is not an uncommon occurrence, as words that have some sort of semantic relationship are often subject to analogy. This is common with numbers e.g., English ‘four’ and ‘eleven’ have been influenced by ‘five’ and ‘ten’, respectively. Balto–Slavic ‘9’ (e.g., OCS devjēť < PIE *H1newn) has likely been influenced by ‘10’ (e.g., OCS desjēť < PIE *dekṃ), though there are other proposed solutions.
generalization looks like it has take place. These cases are likely the result of a number of local generalizations added up over time.

This project focuses exclusively on inflectional verb morphology. Given the nature of morphology, this is an appropriate place to observe this type of phenomenon. Any language has a large number of verbs, and a large number of verbs means that there is a large amount of opportunities to see conjugational patterns. The behavior of speakers when confronted with a verb to conjugate shows they do not necessarily make broad generalizations (and they really are a lot less broad than are traditionally described). While I focus on a specific area of grammar, it is possible to find these “locally based and highly particularized generalizations” in most areas of grammar; see e.g. Gross (1975) for complementation properties for French, Janda & Joseph (1986, 1988, 1992, 2002) for Sanskrit reduplication, Janda (1987) for German umlaut, Stewart (2004) for Celtic Mutations, Neikirk–Shuler (1996) for the adaptation of borrowed verbs into the Slavic morphological aspect distinction, etc.

In what follows I provide several examples of analogical change and demonstrate how the generalizations that the speakers made are rather local, even when in some cases the effects look superficially like an across-the-board generalization.

1.1.2. Paradigm “leveling”

Paradigm leveling is a common phenomenon in which morphophonemic alternations are leveled out within a paradigm. An example can be seen in the present indicative active paradigm of Lat. POSSUM ‘be able’ and its descendants in Spanish and Daco–Romanian.
The Latin verb forms are combinations of a morpheme *pot-* and a conjugated form of *SUM* 'to be'. When the initial sound of the conjugated form of *SUM* was /s/, the final sound of /pot/ assimilated. As three of the six present forms of 'be' had initial s, three of the forms of *POSSUM* had a stem *poss-*. The other three forms, as they did not involve an s-initial form of 'be', did not undergo any changes. However, this stem allomorphy was completely wiped out in Spanish and Daco–Romanian in favor of the forms with the stem consonant /t/ (which regularly voices to d in this position in Spanish).

However, it is not always the case that there is a complete generalization within the paradigm. Sometimes the results are rather more local in nature. For example, in some other Romance languages, the root allomorphy is not completely eliminated.

Whereas in Spanish and Daco–Romanian the Latin stem allomorphy has been completely wiped out, in Italian the distribution of allomorphs is maintained as it was in Latin\(^2\) (though the 2SG and 3SG forms have lost the original /t/). In Portuguese, the stems of

\(^{2}\) Though there is leveling in the infinitive Lat. *POSSE* \(\rightarrow\) It. *potere*.
the 1PL and 3PL have been leveled out, but 1SG *posso remains (rather than an analogized form like *podo).

Another case in which some forms of the paradigm are leveled but others are not can be seen in the development of the present paradigm of Lat. VOLÔ ‘want’ from Indo-European. In Indo-European, reconstructed present systems can be divided morphologically into two groups, athematic verbs and thematic verbs. Besides having some differences in verb endings\(^1\), thematic verbs had the IE thematic suffix *e/o between the root and the endings (e.g., *bher-o-H\(^2\) ‘carry–1SG’), while athematic verbs did not (e.g., *H\(^2\)es-mi ‘be–1SG’). Certain originally PIE athematic verbs have been inherited into Latin with some forms thematized (that is, they gain a theme vowel, and follow the general ‘thematic’ pattern). However, this is not necessarily the case with all the forms in the paradigm. Below is the reconstructed present indicative active paradigm of PIE *wel–, as well as its descendant in Latin.

<table>
<thead>
<tr>
<th></th>
<th>PIE</th>
<th>Latin outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(athematic)</td>
<td>(athematic) (maintained)</td>
</tr>
<tr>
<td>1SG</td>
<td>*wel–mi</td>
<td>→</td>
</tr>
<tr>
<td>2SG</td>
<td>*wel–si</td>
<td>→</td>
</tr>
<tr>
<td>3SG</td>
<td>*wel–ti</td>
<td>→</td>
</tr>
<tr>
<td>1PL</td>
<td>*wel–mos</td>
<td>→</td>
</tr>
<tr>
<td>2PL</td>
<td>*wel–te</td>
<td>→</td>
</tr>
<tr>
<td>3PL</td>
<td>*wel–nti</td>
<td>→</td>
</tr>
</tbody>
</table>

Table 1.3: PIE/Latin PRES.IND paradigm of PIE *wel–/ Lat. volô

The generalization of this originally athematic verb to a thematic pattern of conjugation was not complete, as can be seen in the 3SG and 2PL forms, which maintain the

\(^1\) The main difference between the PIE reconstructed present endings for athematic and thematic verbs is in the 1SG: *–mi for athematic verbs, *–H\(^2\) for thematic verbs, e.g., *H\(^2\)es-mi ‘be–1SG’ (thematic) vs. *bher-o-H\(^2\) ‘carry–1SG’ (thematic). In addition, the 3PL athematic verb ending –enti lacks the initial vowel of the ending in thematic verbs, e.g., *H\(^2\)s–enti (athematic) vs. *bher-o–nti (thematic).

\(^4\) The origin of the 2SG form vîs is unclear. It is potentially from a different root, or perhaps PIE *welsi → *wels → vîs (via some sort of l–vocalization process not attested anywhere else in this environment), Sihler (1995§484).
athematic conjugation. The 2SG form vis makes the evolution of this specific verb even more complicated, as it does not seem to fit into either of the two other patterns. Nevertheless, the generalization was restricted to a few forms in the paradigm, and there is no real reason why the generalization did not happen completely, or why it happened in the first place. Definitely, examples abound of stem alternations being leveled out within a paradigm, thus showing a generalization across the entire paradigm. However, in this case, there has been a generalization, but is has been limited to a certain set of forms. A similar situation can be seen in a few other Latin verbs, e.g., ferō ‘carry’, which has certain forms thematized (actually the same set as the paradigm of volō, except that the 2SG of ferō is fers, which is clearly athematic). Still, this is a very local generalization, restricted both in lexical space as well as within the paradigm.

1.1.3. Generalizations across paradigms

An example of a limited generalization across paradigms can be seen in the development of a small group of verbs in Spanish. Standard Modern Spanish has a group of four verbs that show a 1SG.PRES.IND ending in -oy. In each of these cases, this is not the expected outcome from Latin, as the 1SG.PRES.IND.ACT ending was ō, (and in a few cases, -m). Information from Old Spanish texts show that accretion of this –γ is relatively recent. Below are the Modern Spanish verbs showing 1SG.PRES.IND –γ, along with their 13th century Spanish and Latin descendants:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SUM</td>
<td>so</td>
<td>soy</td>
<td>‘be–1SG.PRES.IND’</td>
</tr>
<tr>
<td>dō</td>
<td>do</td>
<td>doy</td>
<td>‘give–1SG.PRES.IND’</td>
</tr>
<tr>
<td>vādō</td>
<td>vo</td>
<td>voy</td>
<td>‘go–1SG.PRES.IND’</td>
</tr>
<tr>
<td>stō</td>
<td>estō</td>
<td>estoy</td>
<td>‘be–1SG.PRES.IND’</td>
</tr>
</tbody>
</table>

Table 1.4: Four verbs with unexpected 1SG –γ (Wanner 2006)
There have been several accounts for this development; an earlier accepted account is that this final –y (phonetically [j]) is from the Medieval Spanish locative pronoun y (cf. Cat. hi, It. ci), potentially also seen in Sp. hay ‘there is/are’ (Lloyd 1987). However, Wanner (2006) demonstrates that the sequences of these verbs plus y are not found in the corpus of Old Spanish texts, thus putting this solution into doubt. Wanner (2006) proposes this development began with the 1SG of ‘be’, so, becoming soy via influence from the 1SG preterite of the same verb, fui. Then this one case influenced earlier do, vo, estó to become doy, voy, estoy respectively. This solution is advantageous as it explains the significantly earlier attestation of soy over the other three verbs that were affected by the subsequent spread of –y.

Under Wanner’s account of this phenomenon, this is a definitely seen as a local generalization, all beginning with one single form so undergoing a change under the influence of another form, fui. Then, from this stage, it was spread to one of the other verbs (the specific verb actually differs depending on the dialect area, e.g., do in the North, esto in Central Spain). From there it spread to the two other verbs, and precisely there is where the spread stopped. Definitely, there are unifying phonological characteristics of these verbs, but this still does offer explanation into why this change happened and why it was necessarily limited to these four verbs.

Another interesting fact about this change has to do with the verb ‘go’. The 1SG.PERF. of ‘go’ is also fui. However, soy is easily the earliest attested of the four verbs with –y, while voy comes along rather late. So, voy was just as eligible a candidate to have been the starting point for this localized change. It eventually reached this verb, but did not start there.
1.1.4. Extension of 1SG.PRES.IND.ACT -m in West & South Slavic

In some cases, a change may look like a widespread generalization has taken place on the surface. However, upon closer examination, this is not necessarily the case. As discussed above, reconstructed Indo-European verbs were either thematic or athematic. Slavic athematic verbs have a 1SG.PRES. ending from the PIE primary athematic ending *-mi, while the thematic verbs have an ending from *-H₂ (the PIE primary thematic ending), plus *-m (secondary ending). The athematic ending became -мь by OCS, while in the thematic ending the laryngeal is deleted (leaving behind lengthening on the previous vowel) and -m is deleted (leaving behind the nasalization of the previous vowel), eventually universally giving -о. The populations of these two main groups were very lopsided, as in OCS there were only five “athematic” verbs, (i.e., only five verbs had a 1SG.PRES. in -мь,) while all other verbs had a 1SG.PRES. in -о:

<table>
<thead>
<tr>
<th>ATHEMATIC</th>
<th>THEMATIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.PRES</td>
<td>1SG.PRES</td>
</tr>
<tr>
<td>дать 'give'</td>
<td>always in -о, e.g.,</td>
</tr>
<tr>
<td>есть 'eat'</td>
<td>nesо 'carry'</td>
</tr>
<tr>
<td>есть 'be'</td>
<td>delо 'do'</td>
</tr>
<tr>
<td>вёмь 'know'</td>
<td>slyо 'hear'</td>
</tr>
<tr>
<td>имать 'have'</td>
<td>milо 'pity'</td>
</tr>
</tbody>
</table>

Table 1.5: Athematic/Thematic OCS 1SG.PRES.

As for the development of verbs into Modern Slavic, East Slavic is rather conservative. While some of the verbs maintain the OCS athematic conjugation (e.g., Rus. damь 'give-1SG.PRES', emь 'eat-1SG.PRES') there are no additions to this class and OCS имать, one of the five original members of this group, now follows the predominant pattern in East Slavic, e.g., Bel. majу, Ukr. majу, Rus. imeju).

5 There is a variant 1SG form of this verb: вёлешь showing the outcome of middle endings (the only form in which reflexes of PIE middle endings show up in Slavic) (Lunt 2001).
However, this situation drastically changed in South and West Slavic, where this 1SG.PRES.IND. ending \(-m\) that originally had a very limited distribution was extended to other 1SG.PRES.IND. forms. In Bulgarian and Czech it has been extended to some verbs (e.g., Bg. \(iskam\) ‘want–1SG’) while not others (e.g., Bg. \(četa\) ‘read–1SG’). In Standard BCS, it has been extended further, namely to all verbs except for two, \(mogu\) ‘be able–1SG.PRES.’ and \(hoću\) ‘want–1SG.PRES.’.

In Slovak and Slovene it has been extended to all verbs (e.g., Slovene \(hočem\) ‘want–1SG’ cf. OCS \(hošt\); Slovak \(môžem\) ‘can’, cf. OCS \(mog\)). However, there is a consistent theme vowel throughout the present paradigm (even in the 1SG), e.g., Slovene \(hočem\) ‘want–1SG’, \(hočič\) ‘want–2SG’, \(hoče\) ‘want–3SG’, etc. In Macedonian, \(-am\) has been extended to all verbs, regardless of the verb’s theme vowel (e.g., \(možam\) ‘be able–1SG.PRES.’, cf. 3SG \(može\); \(cit\) ‘read–1SG.PRES.’, cf. 3SG \(cít\); \(rabotam\) ‘work–1SG.PRES.’, cf. 3SG \(rabot\)).

In addition to the difference in the 1SG, there was also a difference in the OCS 2SG present forms of athematic verbs and thematic verbs. Unlike the distinction in the 1SG where the thematic and athematic endings have different origins, the two variants seen in the 2SG, \(-si\) and \(-ši\), have the same origin. Both come from earlier \(*-se\)\(^6\), which became \(-ši\) when following a high vowel\(^7\), creating phonologically predictable allomorphy. The palatalized variant was then extended to all verbs except the athematics. The only difference between the distribution of 2SG \(-si\) / \(-ši\) and 1SG \(-m\) / \(-o\) is that ‘have’ has an athematic 1SG and a quasi–thematic 2SG. Examples below:

---

\(^6\) From the reconstructed PIE ending \(*-si\), we would expect Slavic \(-s\), because the PIE ending was short. However, the OCS ending is clearly either from a long vowel or a diphthong, and interestingly, the vowel in the OCS ending was lost throughout Slavic (except in the cases where the athematic ending was conserved!)

\(^7\) When in the RUKI environment (see Andersen 1968), \(*s\) became \(x\). Then, as \(x\) was before a front vowel, it was palatalized to \(š\) via the 1\(^\text{st}\) Slavic Palatalization.
There is no definitive explanation as to why this 2SG -ši was extended to all of the other verbs except for the athematics (it could be mirroring the 1SG in being “different”).

What happened to 2SG of ‘have’, though? In OCS, while the 1SG is imamь, the 2SG is imaši, rather than imasi. The interesting point here is that while 1SG -m has been extended to additional verbs in West and South Slavic (and every single verb in some of these languages), 2SG -si has not been extended whatsoever. It appears that the final vowel of the thematic ending -ši was reinterpreted as a short vowel ŗ, which later was elided. Actually, in West and South Slavic, the 2SG in -š has been generalized to the descendants of the OCS athematic verbs as well. Below are the modern West/South Slavic equivalents of OCS dasi ‘give-2SG’ and ěsi ‘eat-2SG’:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>dasi</td>
<td>dadeš</td>
<td>dadeš</td>
<td>daš</td>
<td>daš</td>
<td>daš</td>
<td>daš</td>
<td>dasz</td>
</tr>
<tr>
<td>ěsi</td>
<td>jadeš</td>
<td>jadeš</td>
<td>jedeš</td>
<td>ješ</td>
<td>jíš</td>
<td>ješ</td>
<td>jesz</td>
</tr>
</tbody>
</table>

Table 1.7: 2SG.PRES. forms of ‘give’ and ‘eat’ in S. & W. Slavic

While the majority of the earlier ‘athematic’ generalized their 2SG to the thematic pattern, the same is not true for ‘be’:
Table 1.8: 2SG.PRES. forms of 'be' in S. & W. Slavic

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>esi</td>
<td>si</td>
<td>si</td>
<td>jesi/si</td>
<td>si</td>
<td>jsi</td>
<td>si</td>
<td>jesteš³</td>
</tr>
</tbody>
</table>

The only other cases of a 2SG in –*si* are found in other athematic verbs that have been maintained in much of East Slavic. Belorussian and Ukrainian maintain the earlier 2SG athematic forms completely, e.g., Bel. dasi 'give–2SG', jasi 'eat–2SG'; Ukr. dasi 'give–2SG', jisi 'eat–2SG' (Mayo 1993, Shevelov 1993). While most dialects of Russian, including the standard, have the equivalent outcome in the 2SG of that of West/South Slavic (i.e. it is analogized/palatalized, e.g., Rus. daš’ 'give–2SG', eš’ 'eat–2SG'), other dialects maintain the athematic 2SG forms (e.g., Northern dasí, 'give–2SG', esí 'eat–2SG', Kasatkin 1989⁹). However, what is it that singled 'be' out as being so different? There are other oddities about this verb in these languages (e.g., suppletive perfective).

However, what actually stood in and prevented the form in, say, Macedonian from being, say, *ši* instead of attested *si*? There is no obvious answer.

On the surface, this change looks fairly general. It covers certain classes of verbs fully in Czech and Bulgarian, all but two 'irregular' verbs in BCS, and all verbs in Slovak, Slovene and Macedonian. However, there are reasons why the generalizations made in this process can be seen as more 'local' than widespread. First, this probably did not all happen at once (see Janda 1996); there was not something that switched on so that suddenly all verbs/classes of verbs had 1SG –*m*. The process of extending this suffix progressed slowly through the verbal system. Also, this can be seen as limited in scope, as 1SG –*m* was not extended to other (non–present) 1SG forms. There is probably was some sort of special characteristic belonging to these verbs (perhaps their

³ The outcome of PIE *s* following a front vowel, having undergone RUKI and the first palatalization in Polish is orthographic <sz>. Orthographic <ś> comes from original palatalized s. The verb has other complications that need not be explained here.

⁹ Even more interestingly, Kasatkin (1989) also reports forms such as dasiš ‘give–2SG’ and esiš ‘eat–2SG’, showing the inherited forms in –*si* undergoing analogy to follow the predominant pattern by the suffixation of –*ś*, rather than replacing earlier –*si* with –*š*. 
high frequency) that led them to resist analogical pressures and remain athematic in OCS. However, why did this irregularity get spread, when the regular form eliminating the irregular form is more common (at least when looking at the trends of analogical change). Even if it is accepted that this small group of verbs is exceptional in some way, then why did its distinctive 1SG ending spread, when its nearly-as-distinctive 2SG ending did not, and eventually fall completely out of use except in the Pan-Slavic outcomes of ‘be’, as well as in a few of the athematic verbs in East Slavic? Definitely, reinterpreting –ṣi as š can be much easier to motivate phonologically than changing original –ǫ into –ь. However, there is no sound change that takes care of this for us. In this change, as is most analogical change, here is no way to surely predict if a change will happen, when it is going to happen and what direction is it going to take (Kuryłłowicz 1947).

The preceding examples shows that analogical change is not something that acts in a wide-spread manner. Even in cases where it looks like that there was a large generalization, often times there is evidence that the spread of this change was gradual, and occurred item-by-item.

1.2. ON INFLECTIONAL CLASSES

1.2.1. Morphology, morphologization, etc.

I firmly argue that the domain of morphology exists, as opposed to some theories that claim that everything falls into phonology or syntax. In addition, it is not unfair to assume that the domain of morphology is more expansive than that of phonology or syntax. A window into this view is shown by the vast numbers of cases of diachronic cases of morphologization from syntax and phonology, while cases in the other direction do exist (from morphology to phonology or syntax), but they are rarer (Joseph & Janda 1988).
There are numerous examples of something originally pertaining to the domain of syntax evolving into something morphological. For example, Lat. MĒNS (GEN.SG. MENTIS), a noun originally meaning ‘mind’, along with being maintained as a noun in Romance (e.g., Cat. ment ‘mind’), also became a suffix (e.g., Cat. –ment) that attaches to adjectives to form adverbs (e.g., Cat. ràpid ‘fast’, ràpidament ‘quickly’). Instances of something phonological becoming something morphological are also quite numerous. As is discussed below, when an originally phonological rule is no longer purely phonological, it moves into the realm of morphology (i.e., morphologically conditioned phonological rules are really morphological rules, Janda 1987). For example, the Celtic mutations (nasalization and lenition) were originally triggered by a originally present final nasal in the case of nasalization and an originally present final vowel in the case of lenition. Due to sound change, the triggering sounds were often lost. However, the effects that they had (namely, leniting and nasalizing) remained, and the result is that some words lenite the first consonant of the following word and some words nasalize the first consonant of the following word. For example, Old Irish a (< *esyo), ‘his’ lenites the first consonant of the following word, e.g., bó ‘cow’, a bó (pronounced [avo]) ‘his cow’. However, along with the lenitng pronoun a, there is also a nasalizing pronoun a which means ‘their’ (< *eisōm) (Thurneysen 1964), e.g., bó ‘cow’, a mbó ‘their cow’ (Quin 1975). This clearly shows that the mere phonological structure of the preposition, synchronically speaking, clearly has nothing do with the mutations.

The question that must be asked is if speakers derive these forms. Is there an underlying representation in which these segments are present (e.g., an underlying nasal in words that have the property of nasalization in Old Irish) that is later subject to phonological rules that always lead to their deletion (or modification)? Or are these morphological rules that are associated with the various words? The second option seems more realistic, as it avoids the proposal of abstract underlying representations
containing elements that are never realized on the surface. Other processes of morphologization often discussed in the literature (though there are still attempts at accounting for them in purely phonological terms) include the various repercussions of Grassmann’s law in Sanskrit (Sag 1974, 1976, Schindler 1976, Janda & Joseph 1985, 1989, 2002).

In cases such as German Umlaut and Sanskrit reduplication, originally simple patterns that covered a large set of forms have fractured over time into a number of patterns, each with a limited lexical range. The development of German Umlaut, from a single phonological rule to dozens of morphological rules, is arguably the best-known example of this phenomenon. In Old High German, an originally phonological rule fronted (umlauted) vowels when the vowel in the next syllable was followed by [i] (although it appears that there is controversy as to whether the rule was already morphologized at this point). This is not an uncommon type of assimilatory process (often also called metaphony or vowel harmony), and it is quite phonetically natural. In certain paradigms, this created stem allomorphy, as some forms within the paradigm would undergo the process while others would not (depending on whether they fit the conditioning environment or not, i.e., if there was an i in the suffix). For example, in the Old High German noun gast, the nominative singular would not be affected (no suffix), but the nominative plural gesti was affected, as the NOM.PL. marker was –i (Wurzel 1981:449). Over time, analogy occurred in certain paradigms (such as that of gast) where the non-umlauted stem was generalized to the singular forms, and the umlauted stem was generalized to the plural forms. At a later time, all short unstressed vowels were reduced to [a], which obscured the environments in which the phonological rule originally applied (as unstressed i would become a), as well as obscuring the phonetic motivation for the rule/process. So, phonological rules could not account for the facts anymore, in the absence of morphological conditioning. These actually became
morphological rules, where umlaut is associated with various stems and various
categories, and widespread generalizations are not possible to make. It is claimed that
there are up to sixty (Joseph & Janda 1989, though I have also heard this number go as
high as eighty) ‘umlaut rules’ in Modern German. Any purely phonological analysis to
account for this data will be ridiculously complex and have to resort to having
morphological features in the conditioning environments (so that it is not ‘purely
phonological’ anymore). In addition, in order to account for cases of umlaut that do not
occur before [i], we would be forced to posit numerous underlying cases of [i] in order to
account for the facts, and since the situation is so complex, the placement of these
underlying cases of [i] would have to be quite arbitrary. The better solution is to
propose a number of morphological rules that are limited in their application. This is
not an uncommon situation:

Unmistakable historical trends involving increased splintering of already-
morphologized processes bespeak a tendency on the part of speakers to opt, in
their internalized grammars, for analyses which focus on individual
morphological and lexical elements, rather than alternative analyses which are
generalized over broader, less idiosyncratic classes of grammatical elements
(whether phonological or syntactic).


While the situation in inflectional verbal systems is somewhat different, there has been
‘fragmentation’ over time, but in the development of Latin from Indo–European (see
chapter 2), as well as from Latin to its various outcomes in Romance (see Chapters
3,4,5,6). While what the factors responsible for such fragmentation are often different
(e.g., the adoption stem extensions in Romance), the principles remain the same.

In describing data such as that concerning German Umlaut and Sanskrit
Reduplication, Janda & Joseph (1986) propose the construct of the “rule constellation”:
This construct (the rule constellation) can be defined as a group of formally similar morphological processes sharing at least one characteristic property of form but distinguished by individual formal idiosyncrasies which prevent their being collapsed with one another.

(Janda & Joseph 1986:85)

They illustrate this with data from Sanskrit reduplication where there are a number of reduplication process that cannot be collapsed into a single “reduplication process”, as they have different features in terms of reduplicant shape, placement (as a prefix or, rarely, word-internally) and various idiosyncratic cases. However, all the rules do share some properties, which point to some sort of ‘unity’ (Janda & Joseph 1986, Joseph & Janda 1988).

While proposed for situations such as Sanskrit reduplication, the constellational model has used in describing other phenomena, e.g., the unity of the Greek negators μη (νι) (Janda & Joseph 1999), mutations in Scottish Gaelic (Stewart 2004), the various forms of a Greek “unceremonious term of address” (Joseph 1997), among others. The notion of “rule constellation” is used in the description of ‘inflectional classes’ in this dissertation. A more specific discussion on how this applies to my view on inflectional classes is discussed in the following sections.

1.2.2. Inflectional classes

There is some ambiguity as to what constitutes an inflectional class; for instance, Dressler (2003) writes that “generally no distinction is made whether a class consists of morphologically identical paradigms or just similar ones”. Some scholars (including traditional sources) use the notion of inflectional class to describe items that inflect in a similar pattern, while some use this to describe a group of lexical items that inflect

---

10 While the most common type is CV reduplication, there are are several other reduplicant shapes, e.g., V, VC, CVC, CVCV, etc.

11 Most frequently πέ or βπε in Greek, but Joseph (1997) lists 56 different forms. This is found elsewhere in the Balkans as well. For more on its use in Greek, see Costanzo (2009).
exactly the same. I take the latter approach, which is discussed in the following section
discussing the general framework in which these data are to be analyzed.

However, instead of using the terminology of “inflectional class”, I prefer to refer
to these as “inflectional patterns”. Two verbs follow the same inflectional pattern if and
only if they follow exactly same pattern (rather than two similar patterns, or two
patterns that share every feature except for one). This definition is precise, and aims to
get away from the ambiguous nature of the definition of ‘inflectional classes’. In
addition, the terms ‘inflectional class’ and ‘conjugational class’ have a connotation with
traditional accounts of the languages to be discussed here (e.g., the four Latin
conjugational classes, the ten Sanskrit classes, etc.).

The ambiguous nature of what defines a verbal class is nothing new. Pāṇini’s
Aṣṭādhyāyī is a list of nearly four thousand statements (or rules) that serves as the
earliest Sanskrit grammar (and the earliest generative grammar of any sort). In addition
to the rules, there were auxiliary texts containing other information. Relevant here is
the Dhātupāṭha, essentially a list of Sanskrit verbal roots, divided into classes. The
classes are based on how roots form their present stem, and each root is classified into
one of ten classes. This classification, while a precursor of modern studies on
inflectional classes, can be seen as problematic here, as there are complicating factors.
The most serious issue is that the aorist stem of a root is not predictable from the
present stem (so, this does not take into account aorist stems). In addition, none of
Pāṇini’s ten present classes is homogeneous and there substantial variation in each.
Similar trends are in early descriptions of Latin, where “conjugations” are repeatedly
referred to by almost all analysts. For example, from the Ars minor of Donatus:
“How many verbal conjugations are there? Three. What are they? First second third. What is the first? It has in the indicative mode, present tense, singular number, second person in the active and neuter verb, long “a” before the last letter, in the passive, common and deponent, before the last syllable, as *amo, amas, amor, amaris*...”

(Chase 1926: 39)

Other ancient sources mention “conjugations” as well when describing the Latin verbal system, e.g., Priscian's *Institutiones grammaticae* mentions four Latin conjugations. However, as is shown in Chapter 2, each Latin conjugation is not homogeneous. Again, that is not to say that there is no pattern: there are actually lots of patterns, but most of them are limited in scope, covering just a few verbs rather than a large group.

1.2.3. Structure of Inflectional Classes/Patterns

It has just been assumed that classes exist, and I agree (though I refer to them as “patterns”). However, if it is agreed upon that they do exist, under what principles are they organized? Proposals of conjugational class systems in Latin, Romance, among other languages of other languages families frequently make reference to the notion of “subclasses”. However, what is a “subclass” and how does it differ in status from a “class”? In what follows, I describe the framework discussed in Dressler (2003), which aims to deal with the organization of inflectional classes, along with issues such as productivity, etc.

1.2.3.1. Dressler’s framework

Dressler makes a distinction between what he calls “macroclasses” and “microclasses”. Dressler’s microclasses are similar to our definition of “inflectional pattern” described above, and in his words, consist of “the set of all paradigms which share exactly the same morphological and morphophonological generalizations”. In addition,
phonological processes that are “regular” (i.e., completely phonological and applicable elsewhere in the language without respect to morphological conditioning) are not sufficient to propose a new microclass. Thus, just because there is a certain verb shows allomorphy does not necessarily mean that this verb warrants the creation of a new microclass, separate from another microclass follows the same pattern, but without allomorphy. However, this is only in the case of a regular phonological process being involved (this is discussed in much greater detail in the following sections of this chapter).

Unlike “microclasses”, Dressler’s notion of “macroclasses” are defined as not requiring that their constituents inflect precisely in same way throughout, and thus, they are defined much less rigidly (similar to the traditional, ambiguous nature of “conjugational classes”). Microclasses that have a certain feature in common are classified together into a macroclass, and each of the microclasses must inherit a property of the superordinate class. In addition, Aguirre & Dressler (2006), introduce the concept of “mini-microclasses”, which essentially consist of derivatives of a single verb that share all the same properties while having some characteristic(s) to the exclusion of other microclasses, e.g., the only verbs that conjugate in precisely the same manner as Sp. hacer ‘make’ are derivatives of hacer, e.g., rehacer, deshacer, etc.). These are to be distinguished from “isolated paradigms”, verbs that inflect in a precise way that is not seen in any other verb. Dressler claims that these do not form a class, saying that “microclasses, but not isolated paradigms, are productive”.

Dressler (2003) applies to this to a portion of Latin nominal and verbal morphology, and Aguirre & Dressler (2006) do the same, in a much more comprehensive way, for Spanish verbal morphology. Both of these analyses claim that the number of microclasses is quite larger than the number of conjugational classes traditionally considered for that language. For example, Spanish verbs are traditionally
considered to be divided into three conjugations (of course, this does not amount to a claim that there are only three ways of conjugating verbs in Spanish). Some verbs are traditionally called irregular, but they are usually described as something like “irregular verbs of the X conjugation”. While Aguirre & Dressler only propose two Spanish macroclasses, they do propose 27 microclasses (a few of which are considered mini-microclasses). In addition, they also propose 22 isolated paradigms (which they claim are not classes, but still, as is discussed below, represent patterns\textsuperscript{12}). Adding these two gives 49 “patterns”. Some of these patterns share most of their features, but even one difference is enough to warrant them being classified as separate.

1.2.3.2. Issues with Dressler's framework

Dressler’s framework is rather helpful, and I completely agree with several of its points. However, there are a few features of this framework that I find problematic.

First, this framework relies on microclasses branching out from a superordinate class. However, is there any psychological reality to this? Speakers definitely realize when patterns share a certain characteristic (as in Joseph & Janda’s notion of the rule constellation). However, which features are relevant to speakers in recognizing these similarities? Assume that two patterns, Pattern A and Pattern B share a feature X warrants them branching from the same node. Now, also assume that Pattern A and Pattern C (and not Pattern B) share a feature Z that warrants them branching from the same node. In this case, which two patterns should be classified together? Does it depend on the salience of the two features, and the two that share the more salient feature get classified together? How do we know which feature is more salient? This is seen in analyses of Romance verbs quite frequently. For instance, under this framework, the Spanish verbs tener ‘have’ and venir ‘come’ to be isolated paradigms within a

\textsuperscript{12} Even if just one verb shows a specific pattern, this can also be seen as a generalization, as speakers have generalized over the various instances of this verb (Joseph 1997a).
macroclass made up of Spanish verbs in –er and –ir. This macroclass is immediately is
split off into –er and –ir ‘classes’ (basically defined as “mini–macroclasses” whose
branching from the same node accounts for the similarities that verbs in –er and verbs
in –ir share). However, tener and venir have several features in common to the
exclusion of most other verbs in their respective traditional conjugational classes,
including non–etymological g in the same distribution\(^\text{13}\), e.g., tengo ‘have–
1SG.PRES.IND’, vengo ‘come–1SG.PRES.IND’. Is this feature (among others) not salient
enough for speakers to not use them in verb categorization? Similar to the traditional
accounts, Aguirre & Dressler’s account considers the theme vowel the be “above” all
other features; as long as the theme vowel is different, there are no further connections
between these two verbs.

In addition, these branching diagrams imply strict boundaries between patterns.
While this might be sufficient if one is concerned with a standard language, where there
is typically one and only one way of conjugating a verb; elsewhere there is an immense
amount of variation, both between speakers and within individual speakers. Verbs
changing their inflectional pattern is not just something that has happened historically;
it happens in everyday speech. Perhaps movement is possible within microclasses of a
single macroclass. However, this is definitely not always the case, e.g., in most Balkan
Romance languages, there is very frequent movement across what would undoubtedly
considered in this framework to be two different macroclasses (as discussed in Chapter
4).

Another issue with this framework is Dressler’s insistence that “isolated
paradigms” are not considered to be (mini–)micro–classes because they are “not
productive”. However, it is possible for a pattern that was only followed by one item to
attract other followers. For example, while the expected NOM.SG. form of the noun

\(^{13}\) This velar consonant is discussed in much greater detail in Chapter 5.
'month' would be *meís* in the ancient Elean Greek dialect, the actual occurring form was *meús* (cf. Attic Grk. *μήν*). The source of this innovative nominative singular is analogy to the NOM.SG. of *Zeús* ‘Zeus’. The interesting point here is that *Zeús* is the only noun (before the analogy) that had this pattern (Anttila 1972). So, this is a case of a “isolated paradigm” being, in a way, productive. This can also be seen in one of the examples discussed earlier, namely, the accretion of –*γ* onto a limited number of Spanish 1SG.PRES.IND.ACT. forms. Textual evidence shows us that this started in one verb, then it spread to the other three verbs, one at a time.

This is not to say that this framework is useless; it is, in my opinion, excellent but flawed. In the following section, an alternative method for defining conjugational patterns and modeling their organization is discussed, borrowing certain aspects from Joseph & Janda (1986, etc.), Dressler (2003), among others.

1.3. A DIFFERENT TAKE ON CONJUGATIONAL PATTERNS

1.3.1. Defining different conjugational patterns

Here I take inspiration from Maurice Gross and his 1979 paper “On the failure of generative grammar”. In trying to construct a generative grammar for French, Gross and his associates put together tables showing the complementation characteristics of verbs – information that would have to be known in order to generate all and only the appropriate grammatical sentences of French. However, what he found is that when all the facts are taken into account, no two verbs show exactly the same properties. He concludes that this information is not generalizable over many lexical items, but is just a property of each lexical item. This goes against some of the prominent ideas of the time, and is consistent with the view that the generalizations speakers make are actually quite more limited in scope than those that most linguists make.
A similar method is employed here; however, instead of exhaustively listing all the complementation properties of verbs, I exhaustively consider all of the different conjugational patterns in the languages in question. I attempt to move away from the current conception of “conjugational class”, and move towards a recognition of different patterns, as many as are needed to list exhaustively all and only the appropriate verbal forms of the language.

This approach is intended to show the whole wide range of conjugational patterns in these languages, and how boiling them down to a small number is actually obscuring the reality of the situation. In honor of Gross, I term this method the “Gross Detail Method” (henceforth “GDM”). The basic tenets of this method are given in the following sections.

1.3.1.1. All details matter

Every single detail, no matter how small, is important. Nevertheless, linguists often gloss over small (or not-so-small) details. For example, Williams (1981) uses Latin data in discussing his “Right-Hand Head Rule”. However, when discussing Latin nominal morphology and his “theory of the paradigm”, Williams leaves out much of the nominal inflectional system. For instance, he only covers four (of the six) Latin cases and three (of the five or more, depending on one’s analysis) declensions (Joseph & Wallace 1984). Why is this data ignored? Because it would complicate the theory Williams proposes. On the surface, an account that looks clean and simple may look elegant. However, elegance at the expense of fully accounting for the data is just superficial.

Any theory has to account for all of the data, and so-called “problematic” aspects should not be discarded. If speakers pay attention to small details, then those who are describing and trying to account for phenomena in these languages should pay attention to these details as well, especially if, as advocates of generative grammar often
say, they are trying to mirror the native speaker’s knowledge of their language. So, if there is a small difference, e.g., an unexpected long vowel where you would expect a short vowel, linguists should not ignore this. If speakers of said language have a long vowel in this environment, they themselves clearly do not ignore it and it is not “unexpected” to them. All of the details matter to the speaker, and if it matters to the speaker, it should matter to the linguist. In fact, it can be argued that these should be of more interest to linguists, as it is in these small and often “peripheral” cases that are the most indicative of what is actually occurring in the speaker’s mind.

1.3.1.2. Distinguishing “patterns”

The second tenet of this process has to with determining when different lexical items are grouped together under a single pattern and when they are not. Following Gross (as well as Dressler’s notion of “microclass”), when two verbs show any differences in conjugation that are not attributable to purely phonological factors (to be discussed below), they are be considered to follow two separate patterns. The term “pattern” is used instead of ‘conjugational class’ here to distance the view here from traditional views that ignore details and group verbs together based on vague similarities and historical reasons.

Any difference between the inflection of two verbs results in these being two separate patterns. There are cases where there is no difference in the inflection of two verbs, e.g., Latin PORTō ‘carry’ and LAUDō ‘praise:
There are no differences in the way that these two verbs are conjugated, and thus, they are considered to follow a single pattern. In other cases, some forms can look similar, while others are different. Take PORTŌ ‘carry’ again, along with DŌ ‘give’. These two verbs have several forms that look similar, e.g., 1SG.PRES.IND.ACT DŌ/PORTŌ. However, the 1PL of ‘carry’ is PORTĀMUS, while that of ‘give’ is DAMUS. This comparison shows a difference, as the 1PL of ‘carry’ has a long vowel and the 1PL of ‘give’ has a short vowel. This is enough to consider these two verbs to follow different pattern (though there are additional differences between the conjugation of PORTŌ and DŌ).

This is not to say that similarity is to be ignored, as judging similarity is the basis of categorization and organization. However, it does not, under the view proposed here, have any role in the definition of a pattern.

1.3.1.3. Phonological predictability

As in Dressler’s principles of defining microclasses, it is not the case that any difference in inflection must result in a classification of two items into different patterns. The only exceptions to this are differences that are completely predictable given the phonological rules of the language in question. For example, assume that in Language X, verb A has stem final /k/ and verb B has stem final /n/. The 1SG form in Language X is formed by adding a vowel –i to the stem. Moreover, assume that the resulting surface forms are –
či and –ni, respectively. If palatalizing /k/ to [č] before /i/ is a phonological process in this language that operates without exception (and with no reference to morphological conditioning) and there are no other differences in which verbs A and B are conjugated, then we can assume verbs A and B follow the same pattern. However, if the k → č change is not a purely phonological rule in the language, e.g., if there is a verb C that has stem final –k but a 1SG form ending in –ki, then this is no longer phonologically predictable, and these verbs have to be considered to be following different patterns.

For example, take the phenomena commonly referred to as Grassmann’s law, which states that, in a sequence of PIE aspirates, the first is deaspirated in Indic and Greek.

<table>
<thead>
<tr>
<th>PIE *bhewdh−14</th>
<th>Sanskrit: bodhāti ‘wake’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Greek: πευθομαί ‘ask’</td>
</tr>
<tr>
<td></td>
<td>Old English: bṓdan ‘give’</td>
</tr>
</tbody>
</table>

Table 1.10: Outcomes of PIE diaspirate root in Skt, Grk, OE

In early Sanskrit, this rule was purely phonological, operating without regard to morphological boundaries (Schindler 1976). However, later, phonological conditioning was later lost, and thus, the operation of the rule was no longer purely predictable.

Rather, accounting for the data would require a number of lexically specific rules (Janda & Joseph 1986, 1989, 1991, 2002). Again, just because this rule at one point was easily captured with a purely phonological generalization does not mean that it remained that way. The instant that something loses phonological conditioning, it must be considered to fall under a different pattern. Phonologists have sought to account for these phenomena with a variety of the state-of-the-art techniques of that day’s phonological

---

14 In Sanskrit, dh is the normal outcome of PIE *dh, but b is not the normal outcome of PIE *bh, rather *b, showing that deaspiration has occurred. The same can be seen in the Greek example, the only difference being that IE voiced aspirates become voiceless aspirated stops in Greek. In English, b and d are both reflexes of PIE aspirates, showing that Grassmann’s law did not occur in Germanic.
theory, but none comes near to explaining what is really going on without having to posit ad-hoc underlying forms that are only posited to give the correct outcome.

More specifically to verb conjugation, we often see phenomena/allomorphy that were based on purely phonological rules can become generalized to forms that did not fit the domain of original application. For example, in a subset of Latin past participles, the appearance of the allomorphs –TUS and –SUS was originally governed by a phonological rule. As soon as there was extension to verbs whose phonological form does not contain the specific conditioning environment, then verbs showing this originally phonologically-conditioned allomorphs will have to be considered to be following different patterns.

In addition, this also applies in cases where the original conditioning environment of a phonological change is lost (e.g., as in German umlaut, Celtic mutations, etc.). For example, in the evolution of Eastern Romance, the 2SG.PRES.IND. had an ending –i (e.g., Italian dorm–i ‘sleep–2SG.PRES.IND.’). In Aromanian (as well as some other Balkan Romance dialects), there is a phonological rule than palatalizes consonants before /i/. However, in some cases, this conditioning environment has been deleted, as final cases of *–i have been lost in Balkan Romance. Proto–Balkan Romance *vin–i ‘come–2SG’ results in AR yinj (phonetically [γiɲ]), with no final vowel. To account for this, some phonologists would propose an underlying representation yini, and would posit two rules: one that palatalizes /n/ before /i/ and another that deletes final /i/, and then order the rules so that the first is applied before the second. However, is there any evidence for this underlying representation? Is there really an /i/ there in the speaker’s minds whose only purpose is to palatalize the previous consonant and then get subsequently deleted? This makes sense historically, but not synchronically. Under this

---

15 The allomorph –sus was triggered by the appearance of a stem-final dental consonant. See Chapter 2 for a more in-depth discussion.
approach, since the conditioning environment is now gone, then it seems that this rule is no longer completely phonological and has moved on the realm of morphology.

1.3.1.4. Branching and subclasses

In descriptions of verbal classes (along with other phenomena), it is common to see reference to “subclasses” or a variety of branching diagrams. However, I am taking the position that what is traditionally called “subclass” is nothing more than another class (or, “pattern”). I argue that there is no fundamental difference between a pattern followed by, say, thousands of members and a pattern following by one single verb. The same goes for branching; there is no evidence that complicated series of branching diagrams has any relation to a speaker’s mental representation of verb categorization. Branching implies some sort of proposal of sub–classes, which this framework assumes do not exist. These diagrams definitely do help us, as linguists, to understand the data, but it is not necessarily true that such techniques are utilized by the non–linguist in verb categorization.

For instance, some Spanish verbs have stem allomorphy between a diphthongized stem and non–diphthongized stem, which is always consistent in terms of the paradigm slots it occurs (at least in the standard language). The allomorphy is seen between e and ie, between o and ue, and in one case, between u and ue. However, not all verbs with e or o in the infinitive have a diphthong in the expected paradigm slots. The chart below shows four verbs from the traditional Spanish 1st conjugation (verbs with infinitives in –ar):

<table>
<thead>
<tr>
<th>Verb</th>
<th>Infinitive</th>
<th>Stem Allomorphy</th>
</tr>
</thead>
<tbody>
<tr>
<td>dar</td>
<td>dar</td>
<td>ie</td>
</tr>
<tr>
<td>traer</td>
<td>traer</td>
<td>o</td>
</tr>
<tr>
<td>ver</td>
<td>ver</td>
<td>u</td>
</tr>
<tr>
<td>querer</td>
<td>querer</td>
<td>ue</td>
</tr>
</tbody>
</table>
This is easily explainable historically. The e in the stem of *cenar* and that in the stem of *cerrar* have different origins, namely Latin Ė and Ė, respectively. Latin Ė has developed into Spanish e regardless of stress, while Latin Ė has developed Spanish e in unstressed syllables (e.g., *cerramos*) and the diphthong ie in stressed syllables (e.g., *ciérro*).

However, it is not the case that this is the origin of all of these verbs, as members have fluctuated between these patterns regardless of their origin.

The forms with monophthong–diphthong alternation are normally considered to form subclasses of the Spanish 1st conjugation (in the other conjugations, as well). True, these verbs have the same endings as other (non–alternating) 1st conjugation verbs (e.g., *cenar*). The main reason for classifying them in this was is that they have the same stem vowel in the infinitive. However, this classification is rather arbitrary. Why do the classes have to be based on the infinitive? This is just something carried over from (imperfect) descriptions of Latin. Why not judge classes by the presence or absence of monophthong–diphthong alternation? There are also verbs in the traditional 2nd conjugation (verbs in –er) and 3rd conjugation (verbs in –ir) that show the same phenomena (e.g., *dormir*, 1SG *duermo* ‘sleep’; *mentir*, 1SG *miento* ‘lie’).

The question that needs to be asked is what do two verbs need to have in common to be considered the same class? Under the view discussed here, they have to...
have everything in common (except for features that can be accounted for via regular phonological processes active in the language). We could potentially propose underlying diphthongs in the stems for the verbs above that show diphthongs, and then propose a rule that monophthongizes them in unstressed position. However, doing so leads to more abstractness in the grammar, something that this model is attempting to avoid.

1.3.1.5. Ignoring history

Knowledge of the history of the development of a language is part of the domain of linguistics, yet it is something that an average speaker (i.e., a non-linguist) does not have access to. So, when formulating synchronic analyses of verbal classes (or really any aspect of language), diachronic information must be ignored.

The influence of knowledge of earlier stages of the language on the analyses of linguists can be seen in early generative phonology, for instance, where underlying representations often made reference to earlier forms of words of which no any modern speaker would not have knowledge. For example, Chomsky and Halle (1968), when discussing exceptions to their rule of Trisyllabic Shortening, bring up the example of the English word *nightingale*. Given this word and the proposed rule, the first vowel would be expected to be shortened. To get around this, they propose an underlying representation, which among other things, contains the sound [x]. To be sure, this sound was present in this word in earlier forms of English and is found in some contemporary dialects (e.g., Scots English). However, it has been gone for a long time, and there is no reason for speakers to assume that it is there. This is just a case of altering an underlying representation so the analysis works. This is actually quite dangerous, as it basically allows the proposal of anything in the underlying
representation to make the rules work (see issues with ‘absolute neutralization’ and the

The same can be seen in some analyses of the phenomena relating to
Grassmann’s law. While many IE roots can be reconstructed as having two aspirates,
there are nearly none in Sanskrit\(^\text{16}\) due to the effects of Grassmann’s Law. However,
depending on a number of phonological processes, the aspiration can occur on the first
consonant or on the second consonant (or on neither), see below:

<table>
<thead>
<tr>
<th>Root</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bodh-ati</td>
<td>'know–3SG.PRES.IND.ACT'</td>
</tr>
<tr>
<td>bhot–syati</td>
<td>'know–3SG.FUT.IND.ACT'</td>
</tr>
<tr>
<td>bud–dha</td>
<td>'know–past passive participle'</td>
</tr>
</tbody>
</table>

Table 1.12: Forms of Skt. √budh– ‘know’

Early accounts attempted to get around this by proposing underlying representations
with two aspirates, in this case, /bhodh/. The underlying root–initial aspirate would be
deaspirated in the form bodhati, the underlying root–final aspirate would be deaspirated
in the form bhotyati, and both would be deaspirated in the form budha. However,
since aspiration never surfaces in both places in the same form, this underlying
representation is not motivated by the facts of the language. It is motivated by trying to
make this fit a simple phonological analysis, as well as the history, as the reconstructed
Indo-European root from which this descends does have two aspirates (and there is
actual motivation for this reconstruction given the facts of the Indo-European
languages, Schindler 1976).

As for verbal classes, the average speaker and the historical linguist have two
different points of view. The historical linguist knows where certain verb patterns come
from and thus can group two groups together based on historical reasons. However,

\(^{16}\) There are few, e.g., √dhrāgh– ‘be able’, and √bharbh– ‘injure’ though they are late/potentially
creations of the grammarians (Janda & Joseph 2002)
speakers have no idea of these historical reasons. Several examples are shown in the following sections.

Thus, an essential part of the GDM is that history must be ignored. If we are looking at phenomena purely synchronically, then the linguist cannot resort to historical reasons in explaining what is going on in the speaker’s mental representation. Definitely, most of the phenomena to be discussed here have a (often very clear) historical representation. However, this is something to which speakers just do not have access as they are acquiring their language as children.

1.3.2. Organization of patterns

A full formalization of how I believe that different verb patterns are organized cannot be given here. Nonetheless, the following section gives some ideas as to how I conceive of this working.

1.3.2.1. An illustrative example

Take the English verb *differ* and the related English adjective *different*\(^\text{17}\). In earlier stages of English, these words had mainly the same features. To make this simple, assume there to be four features:

1. SEMANTICS
2. PHONOLOGICAL SHAPE (in general, disregarding the suffix)
3. COMPLEMENTATION
4. PART OF SPEECH

In earlier forms of English, *differ* and *different* agree on feature (1) as they share the same core meaning, feature (2) as they roughly have the same phonological shape and feature (3) as they both took the preposition *from*. The only main difference is in

\(^{17}\) This example was suggested to me by Brian Joseph, as an illustrative case he uses in class lectures.
feature (4), namely that *differ* is a verb and *different* is an adjective. If we represent the properties of these words with circles, the circles would be nearly overlapping:

![Venn Diagram](image)

Figure 1.1: *differ* and *different*, earlier English

In modern American English *differ* and *different* no longer agree on feature (3), as *different than* is the preferred construction over *different from*. This is a feature that *different* shares with comparatives (e.g., “The Dodgers are **better than** the Giants.”). *Differ* and *different* still agree on features (1) and (2), while *different* agrees with comparatives in (3) and (4) (as comparatives take *than* and are adjectives).

![Venn Diagram](image)

Figure 1.2: *differ*, *different*, comparatives, Modern English
However, differ does not completely pattern with comparatives, as *different* cannot occur in comparative correlatives (see Smith 2010), such as the following example:

...the sooner the Dodgers get to the Philadelphia bullpen, the better their chances are. 18

While any comparative could be put in this construction, *different* cannot (it must be preceded by *more*). *Different* is just, well, different. This is a comparatively simple example (that has been simplified!), but it shows the principles of an alternative approach to the categorization of conjugational patterns.

**1.3.2.2. Representing rule constellations**

Recall that Janda & Joseph (1986) define a rule constellation as a group of rules that share some characteristic but are not collapsible. There is always going to be some sort of unifying feature, but the other details may differ. Take Sanskrit reduplication, for instance. Joseph & Janda (1986) describe a few features by which the different reduplication rules may be organized.

1. reduplicant shape
   (e.g., CV, VV, V, CVV, etc.)
2. consonantal specification
   (restrictions on the consonants that can occur in a reduplicant, e.g., restriction on aspirated or velar consonants in the reduplicant)
3. placement of reduplicated syllable
   (prefixed or infixed)
4. root idiosyncrasies
   (e.g., a few roots with initial palatals have “reversion” to velar when reduplicated, e.g., root √cit-, perf. stem ci-kit–)

---


Referring to the 2009 NLCS between the Los Angeles Dodgers and the Philadelphia Phillies. Unfortunately, except for in game 2, the Dodgers were not able to capitalize on Philadelphia’s bullpen struggles and bring the National League championship back to Los Angeles.
Then assume that for any of these features there are a number of specifications (e.g., for reduplicant shape, there would be seven different specifications). Then, when drawing a diagram, where a group of circles overlap, this is an individual reduplication rule. This can be taken even further. For the feature in (1) there could be a group of circles that indicate the different features, and there would be overlap where the reduplication pattern had something in common (there would be an area that all the circles share, as all reduplication rules involve the affixation of a vowel).

Note that I am modeling the facts and offering a possible visualization. I am definitely not saying that the minds of speakers are filled with a series of intersecting circles. However, these are all patterns seen in Sanskrit, and if speakers recognize the similarities between the different patterns, then this model is a way of visualizing this.

1.3.2.3. Representing verb classes

As discussed above, I argue that verb patterns can be seen as constellations. All of the verbal patterns in a language are going to have something in common. Certain verb patterns have more in common with some verb patterns than others. In representing verb classes, it is essential that the boundaries between certain patterns not be impermeable, as verbs can change their pattern over time. In addition, of these boundaries were not impermeable, then there would be no way for verbal borrowings to enter the languages and be adapted to the verbal system. In addition, speakers show variation as to the pattern a certain verb follows; sometimes even a single speaker shows such variation. It is true that verbs that change or vary in the conjugational pattern they follow often change to/vary with a pattern with which it shares some features, while it is rare, but by no means impossible, for a verb to radically change its pattern.
If we claim that inflectional patterns are essentially an intersection of various features, then we can perhaps represent the different patterns as spheres in a multidimensional space. Patterns that have more followers would be represented by larger spheres and patterns with just a few followers (or even a single follower) would be represented by a smaller sphere. These spheres could then be placed in this multidimensional space, where spheres that have many features in common would be located close to one another and spheres that share just a few features would be further away from each other. The important thing here is that, in a way, all spheres are equal in the fact that they exist. Thus, no pattern, no matter how irregular is may seem, is never “outside” of the system. It may be a little bit further away than many of the other patterns, but it is definitely there, categorized based to the features that it shares with other patterns seen in the language.

This also does not imply any subbranching. Spheres cannot exist inside other spheres or intersect. However, this model gets at the similarities that, say, two branches from a single node implies. These are still different spheres, but they are located closer together due to their shared features.

This is not an accurate picture of how the mind works. However, are branching diagrams an accurate picture of how the mind works? Traditional pictures (even that developed by Dressler) of Latin conjugation still are organized by the theme vowel. However, with a branching diagram, once one splits the Latin conjugational system into the traditional four branches (representing the four conjugational classes), one is basically saying that there is nothing more significant that some verb of one branch has in common with a verb of another branch. So, for example, when Aguirre & Dressler split their Spanish i/e ‘macroclass’ into an i-class and an e-class, they are basically implying that there is nothing more significant that certain verbs of the i-class have in common with certain verbs of the e-class. However, this just is not true. Visualizing
these patterns as individual spheres whose proximity is based on similarity gets around this issue, as it allows for recognition of similarity of various features simultaneously.

What about productivity? What happens when new verbs enter the language? They get attracted to a specific “sphere” and follow that pattern. Some spheres are going to have some sort of attractive force. This force could be due to the size of the sphere (so, a pattern that already has a large number of verbs following it might attract new verbs) or some other reason (some patterns are just productive, even if they represent a smaller population of verbs).

Once one looks at all the conjugation patterns in a language, the trend that will emerge is that there are many patterns. That is, in this visualization, there are going to be a lot of spheres. There will be some larger spheres (there are definitely patterns that many verbs follow), but the majority of the spheres are going to be small. This is representative of the local generalizations that speakers make.19

1.4. Sanskrit Verb Classes

The following section will demonstrate the techniques presented in the previous section in describing the Sanskrit verbal system. This will not be an exhaustive analysis of the system. The main function of these examples are to show that, if all of the data are taken into account, the picture of Sanskrit verbal classes that starts to emerge is one that is quite fragmented.

---

19 The model proposed here is similar to Wittgenstein’s Family Resemblance model. See the concluding chapter of this dissertation for more on this connection.
1.4.1. The Sanskrit present system

1.4.1.1. Traditional views

Pāṇinian tradition divides Sanskrit verbs into ten classes, based on how each verb’s root forms its respective present stem. Below are the ten traditional classes, accompanied by representative examples:

<table>
<thead>
<tr>
<th>1</th>
<th>a-class</th>
<th>bhvādīgana</th>
<th>[ROOT (strengthened(^{20})) + a]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e.g., root √bhū ‘be’, present stem: bhav-a-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>root-class</td>
<td>adādīgana</td>
<td>[ROOT]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √i ‘go’, present stem: e- (strong)/i- (weak)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>reduplicating class</td>
<td>juhotyādīgana</td>
<td>[reduplicant + ROOT]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √bhṛ, present stem: bi-bhar- (strong)/bi-bhṛ- (weak)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ya-class</td>
<td>divādīgana</td>
<td>[ROOT + ya]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √nah- ‘bind’, present stem: nahya-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>nu-class</td>
<td>svādīgana</td>
<td>[ROOT + no (strong)/nu (weak)]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √kr-, present stem: kr-no- (strong)/kr-nū- (weak)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>á-class</td>
<td>tudādīgana</td>
<td>[ROOT + á]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √tud-, present stem tudā-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>nasal class</td>
<td>rudhādīgana</td>
<td>[ROOT + {na (strong) / n (weak)}]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √yuj- ‘join’, present stem: yu-na–j– (strong) / yu-nī–j (weak)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>u-class</td>
<td>tanādīgana</td>
<td>[ROOT + u]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √tan– ‘stretch’, present stem tan–u–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>nā-class</td>
<td>kryādīgana</td>
<td>[ROOT + nā (strong)/ni (weak)]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √grabh–, present stem: grbh-nā– (strong)/grbh-nī– (weak)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>aya-class</td>
<td>curādīgana</td>
<td>[ROOT + aya]</td>
</tr>
<tr>
<td></td>
<td>e.g., root: √cint– ‘think’, present stem: cint–aya–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.13: Sanskrit present classes (traditional view)

While the ten-class system is the one proposed by Pāṇinī, Whitney (1896) divides the traditional classes into two major groups (or following Dressler’s terminology, “macroclasses”), one called the “first or non-a-conjugation” (comprising traditional classes 1, 3, 5, 7, 8, and 9) and the other called the “second or a-conjugation” (comprising traditional classes 2, 4, and 6). This division is based on certain

\(^{20}\) “strengthening” means the vowel is strengthened to ‘guna’ grade, essentially meaning that a vowel a is added before the vowel of the root, and then the normal sandhi rules apply, e.g., root √bhū → present stem √bh(a)a → bhava- (via normal sandhi rules). A root vowel a does not undergo strengthening, as “a is its own guna” (Whitney 1896 §235)
characteristics each group shares, e.g., Whitney’s 1st conjugation comprises groups that undergo an accentual shift in paradigmatically related forms, e.g., singular vs. plural, while his 2nd conjugation comprises groups in which the stem ends in –a. Whitney also claims that the traditional 8th class is better considered a subclass of the 5th class (though his reasoning for its demotion to subclass status is mainly based on the fact that the 8th class covers a very limited set of verbs\(^{21}\)). The 10th class (or in Whitney’s words “the so-called tenth class”) is, according to Whitney, not really a class at all, but rather a “secondary/derivative conjugation”, inasmuch as it is populated mainly by denominatives and causatives. However, there are cases of verbs that form a stem in –aya (i.e., part of the traditional 10th class) that are neither denominative nor causative (see example √cint– above).

However, these issues relating to the “merging” of classes run into issues similar to those found in Dressler (2002). That is, what types of characteristics that two patterns share are significant enough to justify the grouping of two classes into a class that encompasses them both? The division into two main classes seems like an analysis done after the fact – more of something that Whitney noticed while studying Sanskrit grammar in the late 19th century, rather than a generalization made by actual speakers of Sanskrit several thousand years earlier\(^ {22}\). Is there any evidence that that there are two macro-classes that get broken down? Maybe it looks that way to the linguist, but it is difficult to tell if speakers followed the same principles of categorization.

\(^{21}\) This is justified historically for the 8th class roots than end in a nasal (e.g., √tan–), but since the –u–/–o– suffix has been extended to a few non-nasal roots, most notably √kar–, it must now be considered a different class, even if a small one.

\(^{22}\) It is true that Pāṇinī was a native speaker of Sanskrit (or a language rather similar to Sanskrit). Thus, does his ten-class system represent psychological reality to a native speaker? I would argue that it does not, as Pāṇinī’s grammar was more intended to teach the language, and was constructed intentionally as such.
1.4.1.2. Exceptions in the 10 present classes

Dividing the present stem types into ten classes (following Pāṇinī and most authorities on Sanskrit) accounts for the majority of present stems. That is, if one knows a verb root and which of the ten traditional present classes it belongs to, chances are that the formation of the present stem along with the conjugation of the various present stem categories will come out correctly.

However, when describing the data, the goal is not to account for just most of the data. Rather, following the GDM, everything, i.e., every detail, must be accounted for. A closer look at Sanskrit present stems reveals that exceptions are widespread, that is, not all members of each traditionally-described present class conjugate in precisely the same way and thus, by the point of view espoused here, actually represent different patterns.

For example, the roots √nam- ‘bend’ and √gam- ‘go’ are both traditionally considered to be part of the first class because they both have present stem formed by the addition of –a to the root. However, the present stems are actually different; √nam has the expected nama- while √gam has the somewhat unexpected gaccha-\(^2^3\). While they do show the similarity of having an –a suffixed to the root in their present stem formation, this quite striking difference shows that there is a different pattern, and thus, by the approach taken here, they must be considered as following different patterns.

This is not the full extent of the verbs traditionally considered to be part of the 1\(^{st}\) class that do not form their stem by simply strengthening the root and adding –a. There are examples of verbs that undergo strengthening when not expected, there are verbs that do not undergo strengthening when expected, there are roots that have lengthening of the root vowel instead of strengthening, and so on. The unification of all these verbs under a single heading is not necessarily misguided, but it belongs more to the realm of

\(^{23}\)-cch- is a reflex of PIE *sḱ-*, an affix whose outcome(s) in Latin/Romance are discussed in detail in Chapter 4.
the analysis of the linguist, rather than that of the speaker. They all do share a characteristic, namely the addition of \(-a\) to the root to form the present stem. However, besides that, it is almost the case that anything goes.

Another example lies in the reduplicating (or 3rd) class, where the form of the reduplicant used in stem formation is not always consistent, e.g., \(\sqrt{dā} \) ‘give’ forms a present stem \(dadā\), with a reduplicant vowel \(a\), while \(\sqrt{gā} \) ‘go’ forms a present stem \(jigā\), with reduplication vowel \(i\). The “default” pattern, based on the majority of Sanskrit reduplicated presents would have the reduplicated vowel be \(u\) if the root vowel is \(u\) (e.g., \(\sqrt{hu} \rightarrow \text{present stem } juho\)) and otherwise \(i\) (regardless of the quality of the root vowel, e.g., \(\sqrt{gā} \rightarrow jigā, \sqrt{bhṛ} \rightarrow bibhar\)). This could be understood diachronically, but synchronically, it is necessary to propose distinct patterns involving reduplication in these cases, and thus, proposing distinct patterns of verb conjugation. Clearly, speakers of Sanskrit did not treat all present stems in the ‘reduplicating class’ in precisely the same manner. Thus, while a notion like ‘reduplicating class’ is something valuable to the linguist, it appears as it did not have much significance for speakers of Sanskrit. Even though verbs like \(\sqrt{dā} \) and \(\sqrt{gā} \) both formed their present stem via reduplication, since the details of reduplication in all of these forms is not a unitary process and cannot efficiently be generated by any phonological solution without positing highly abstract\(^{24}\) underlying forms, e.g., saying that different verbs have different reduplication templates with different material already filled in. Not being able to account for everything in a simple manner synchronically is not a flaw, it is reality, and the effort of linguists to make everything “fit” one pattern is misguided. Also, the practice of positing underlying representations of varying levels of abstraction is essentially saying that there are different patterns.

\(^{24}\) Again, by “abstract” I refer to the positing of material in the underlying representation (that often does not appear on the surface) solely in order to generate the right form.
This, while not exhaustive, shows the different ways in which present stems are constructed. Even amongst groups of verbs traditionally considered to belonging to a certain present class, there are differences found within each class. According the Gross Detail Method, as described above, these would have to be considered different patterns. It definitely does matter that √nam- and √gam- or √dā- and √gā- share some characteristics of their conjugation in common. However, under the framework proposed earlier, one difference signals the proposal of different patterns. Pāṇini actually follows in this manner, paying very close attention to the details. For example, in the Dhātupāṭha, as discussed earlier, a number of subclasses are listed within many of the subclasses (these are cases where the verb forms a present stem in a similar, though not precisely the same way).

1.4.2. The Sanskrit aorist system
1.4.2.1. Traditional view

The present system is not the only portion of Sanskrit grammar in which such differences in stem formation exist. A similar situation can be seen in the different Aorist system classes, seen below (again following Whitney's grammar):
There are seven different patterns used in forming the aorist stem from the root. Some of these share some features. For example, options 4 through 7 above all involve suffixing some sort of s-like consonant to the root. However, the details differ among these four options and are not synchronically phonologically predictable. Thus, these patterns need to be considered separately.

### 1.4.2.2. Exceptions to aorist patterns

Again, just as with the present system, the different aorist classes are not completely uniform, as there are several issues that make the complicate the picture. One issue has to do with the fact that several roots (more than fifty in Vedic, Whitney 1898) have more than one aorist formation (e.g., √dhā- has a root aorist and an s-aorist). In Classical Sanskrit, the root aorist is only used in the active, while the corresponding passives use a sigmatic aorist. Again, as in the reduplicated present, the reduplicated aorist shows some idiosyncrasies. In addition, there are numerous unpredictable lengthenings and shortenings of vowels.

---

Note that I have just referred to this as the “reduplicant”. The actual reduplicated material varies depending on the verbal category as well as within verbal categories as well (see Janda & Joseph 1986, 1989)
1.4.3. “Linkages” between present and aorist stems

The issue here is that there are not strong connections between the present and aorist systems, and thus, just because two verbs follow the same pattern in the present systems, this definitely does not mean that they will follow the same pattern in the aorist systems. The exception with this is that the 10th present class corresponds to a great degree (but not entirely so) with the reduplicated aorist (Whitney 1896§856). Otherwise, a verb belonging to any of present stem types can have a corresponding aorist stem of any of the aorist stem types. The chart below shows the different aorist formations for verbs belonging to the 1st present class.

<table>
<thead>
<tr>
<th>root</th>
<th>present class</th>
<th>present stem</th>
<th>aorist class</th>
<th>aorist stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>√bhū</td>
<td>1</td>
<td>bhava-</td>
<td>1</td>
<td>bhū</td>
</tr>
<tr>
<td>√vṛt</td>
<td></td>
<td>varta-</td>
<td>2</td>
<td>vṛta-</td>
</tr>
<tr>
<td>√kas</td>
<td></td>
<td>kasa-</td>
<td>3</td>
<td>cīkas-</td>
</tr>
<tr>
<td>√nī</td>
<td></td>
<td>naya-</td>
<td>4</td>
<td>nāīs-</td>
</tr>
<tr>
<td>√budh</td>
<td></td>
<td>bodha-</td>
<td>5</td>
<td>bodhiṣ-</td>
</tr>
<tr>
<td>√nam</td>
<td></td>
<td>nama-</td>
<td>6</td>
<td>namsiṣ-</td>
</tr>
<tr>
<td>√mih</td>
<td></td>
<td>meha-</td>
<td>7</td>
<td>mikṣa-</td>
</tr>
</tbody>
</table>

Table 1.15: Present class/aorist class correspondences

The same trends can be seen among verbs that exhibit other present stems. When looking at the present and aorist systems together, one sees that the number of patterns exhibited by Sanskrit verbs becomes much larger. When taking into account the fact that not all verbs commonly classified in the same ‘class’ form their present stem in the same way (recall the example of the present stems for √gam– and √nam– above), the true level of the number of patterns in Sanskrit verbal inflection starts to become clearer.
1.4.4. Other details of Sanskrit verb conjugation

The Sanskrit inflectional system is, clearly, not just divided into the present system and the aorist system. There are many more verbal categories in the language, and while some of them show relatively regular and consistent patterns in conjugations, there are others that definitely do not. For example, as is the case in Latin discussed in the next chapter, the formation the Sanskrit past passive participle follows several different patterns that are generally not predictable based on the root.

In forming the past passive participle, the most common pattern is to add –ta to the weak form of the root, e.g., √bhū– ‘be’, PPP bhūta–. However, the exact forms of the root to which the suffix is added is occasionally unpredictable. For example, some verbs in final –m lose the nasal in the PPP, e.g., √gam– ‘go’, PPP gata–, while others maintain the nasal and then lengthen the root vowel, e.g., √tam– ‘go’, PPP tāṁta–. There are verbs that form their PPP in –ita. While it looks like that this is phonologically motivated in some cases, e.g., those where adding –ta would violate some phonotactic constraint (Whitney 1896§956). However, there are many other verbs that have the PPP in –ita where this is not the case. Besides the suffixes –ta and –ita, there is also a suffix –na used in the formation of past passive participles. There are some tendencies to its distribution, but it is by no means completely predictable from the root.

When the present stem classes, the aorist stem classes, and the different patterns in PPP formation are taken together and the “linkages” between the three are considered (along with all other aspects of Sanskrit verbal morphology not considered in this brief section), the actual number of patterns observed becomes quite large. However, the question that remains is how considering all these ways of conjugating verbs as individual patterns is advantageous over other models.
1.4.5. Classification and categorization of Sanskrit verbs

Any model that considers one feature of Sanskrit verb conjugation as primary over all others is going to definitely be flawed. If one were to say that the present system is primary in the categorization of Sanskrit verbs, then that would imply that in a tree diagram, the highest node would be split into ten branches, each representing one of the ten traditional present classes. Then, as each of the present classes has verbs that “link” to different aorist stem types, then each of these ten branches would have to branch into seven additional branches, each of which would represent the linkage between that present stem type and one of the perfect stem types. However, there is no reason to consider the present stem (or the aorist stem, or the PPP formation pattern) to be primary.

I argue that Sanskrit verbs are organized on the basis of similarity. However, I do not consider similarity to be judged primarily on one factor. All of the patterns seen in Sanskrit verb conjugation are independent, but each pattern shares certain features with other patterns. It is true that some features used in the judgment of similarity and the categorization of the verbal system are probably more salient than others. However, it is not the case that one feature has to be considered primary. Several features, even though they may be weighted differently, are judged simultaneously.

1.5. SUMMARY

This dissertation discusses the scope of generalizations, specifically in the domain of inflectional verb morphology. Chapter 1 provided a background to these issues. A close examination of language data shows us that the generalizations that speakers make are quite restricted in scope.

As for inflectional classes, there is no clear definition. Some use the term “inflectional class” for a group of items that inflect in a similar way, while others use it to
describe a group of items that inflect in the exact same way. Adding to the
complication of the term “inflectional class” is the fact that early grammars (e.g., of
Sanskrit, Latin, etc.) use “conjugational class” or “verb class” in describing the formation
of present stems from roots, completely ignoring the formation of other stems within
the same paradigm. For these reasons, I choose to use the term “inflectional pattern”
to describe the exact patterns that verbs follow in inflection (that is, if two verbs show
any difference in conjugation, they follow two distinct conjugational patterns).

To discuss inflectional classes, I borrow from the work of Dressler, and take
some lessons from the work of Maurice Gross. This is called the “Gross Detail method”,
the major tenets of which are: (i) looking at all of the details, no matter how small or
apparently insignificant, (ii) any difference between two verbs shows that these verbs
have different “conjugational patterns”, (iii) the only exceptions lie in completely
phonologically predictable processes, (iv) the absence of hierarchical structure (or
“subbranching” of any sort) within classes, and (v) not taking into account the diachronic
evolution of these “patterns”. This can be visualized by considering the whole
inflectional system to be a three-dimensional space where different patterns are
represented by non-intersecting spheres where similarity is judged by a multitude of
factors (rather than a main split by one factor considered primary as in tree diagrams).

Following this background, the structure of the rest of this dissertation proceeds
as follows: Chapter 2 discusses issues in Latin conjugational patterns in detail, and
shows how the traditional classification is problematic. Chapter 3 discusses basic
considerations of the development of Romance inflectional systems. Chapter 4 through
Chapter 6 discuss these issues regarding specific phenomena in a variety of Romance
languages, using data from grammars, corpora, dialect atlas and fieldwork. Chapter 7
offers closing remarks.
CHAPTER 2:
PATTERNS IN LATIN VERBAL INFLECTION

2.0. INTRODUCTION

At a first, and somewhat superficial glance, the Romance conjugation class systems do not appear to be particularly “complicated” as laid out in most grammatical descriptions. There is a limited set of “conjugations” (usually between two and five) into which all verbs, save a few “irregular” verbs, fall. One of the reasons for this is that Latin has a similar reputation for not being necessarily complicated, and that over time, the system is considered to have undergone simplification in the development from Latin into Romance. On closer examination, this myth of simplicity turns out to be untrue. The Romance verb system is actually rather complicated, and it is so for two reasons: (1) It descends from a language whose verbal system was not exactly regular itself, and (2) the developments that transformed Latin into the various Romance languages, which in some aspects eliminated some of the Latin irregularity, also led to the development of considerable “irregularity”.

In preparation for the discussion of various Romance verbal systems (and specific phenomena within) presented in Chapters 3 through 6, this chapter describes the Latin conjugational system following the principles of the GDM, as outlined (and briefly demonstrated) in the previous chapter. To be clear, this is not an exhaustive description of the system or an exhaustive listing of every single pattern followed by verbs in Latin. Alternatively, this discussion serves (much as the brief Sanskrit example in the previous chapter) to show how complicated the system really is if we take all of
the details of conjugation are taken into consideration. The goal here is to show that, while there definitely were patterns, the generalizations made by Latin speakers were much more restricted in scope than sometimes assumed. The reputation for simplicity of Latin (and consequently, of Romance as well) is the result of the propagation and the maintenance of the “four-conjugation” system in describing Latin, despite the fact that any scholar or learner of Latin obviously knows that there are not only four distinct patterns that Latin verbs follow in conjugation. This four-conjugation system, as shown below, falls short in many respects.

2.1. AN OVERVIEW OF THE LATIN VERBAL SYSTEM

2.1.1. infectum, perfectum, etc.

Traditionally, the Latin verb system is considered to consist of two systems: the present system (infectum) and the perfect system (perfectum) (Hale & Buck 1903). This can be seen as both a semantic, as well as a morphological, distinction. Semantically, infectum generally refers to forms with imperfective (continuous) aspect and perfectum refers to forms with perfective (completed) aspect. This is a morphological distinction, as the present system and the perfect system are formed by different verbal stems that are usually (e.g., CANTĀ ‘sing–PRES’ ~ CANTĀVĪ ‘sing–PERF’), but not always (e.g., FER– ‘carry–PRES’ ~ TULĪ ‘carry–PERF’), derived from the same root.

However, the fit between the semantic classification and the morphological classification is not precise. This can be seen in the past passive participle, which is semantically perfectum, but is formed from a stem that is neither infectum nor perfectum1, e.g., MORSUM ‘bite–PPP’, cf. MORDEŌ ‘bite–1SG.PRES’, MOMORDĪ ‘bite–1SG.PERF’.

In addition, it has been argued that the present imperative and the gerund can be considered semantically of “neutral” aspect, while morphologically being formed from

---

1 There are correlations between the perfect stem and the participial stem, though they are not one in the same. See section 2.4.
the present stem (DeWandel 1982), e.g., CANTĀ–‘sing–PRES’, CANTANDUM ‘sing–GER’.

Thus, strictly morphologically speaking, the Latin verbal system must be divided into three systems (or verbal stem types): the present system, the perfect system, and the “participial” system. The following table shows the different verbal stem types of Latin and the verbal categories derived from each stem:

<table>
<thead>
<tr>
<th>present stem</th>
<th>perfect stem</th>
<th>participial stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>present indicative/subjunctive</td>
<td>perfect indicative/subjunctive</td>
<td>past perfect participle</td>
</tr>
<tr>
<td>imperfect indicative/subjunctive</td>
<td>future indicative</td>
<td>perfect participial</td>
</tr>
<tr>
<td>future indicative</td>
<td>pluperfect indicative/subjunctive</td>
<td>future participle</td>
</tr>
<tr>
<td>present imperative</td>
<td>present participle</td>
<td>perfect infinitive</td>
</tr>
<tr>
<td>present participle</td>
<td>future perfect indicative</td>
<td></td>
</tr>
<tr>
<td>gerund</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 2.1: Latin verb categorization*

In what follows, the different patterns seen within each system (present, perfect, participial) are discussed. In addition, and perhaps more importantly, the linkages between present stems, perfect stems and participial stems are described in detail. That is, is the perfect stem of a certain verb predictable from its present stem (or vice-versa)? In short, the answer is that it is not. Some strong tendencies are observable, but they always come with exceptions (sometimes few, sometimes many). The emphasis here, following the approach described in Chapter 1, is on the details, and a close examination of such details has important consequences for the view presented here that generalizations speakers make are much more limited in scope than those that linguists often man to account for what speakers do.

---

2 This name is borrowed from Hale & Buck (1903). This definitely does not imply that all Latin participles have parallel forms as a base (e.g., the present participle is formed from the present stem). The “participial” stem has alternatively been called the “3 stem” (Aronoff 1992)

3 Of course, most of these verbal types have morphologically distinct active and passive forms.
2.1.2. Traditional views of Latin conjugation

The traditional descriptions of the Latin conjugational class system are based on the quality of the “theme vowel” in the present stem (Hale & Buck 1903, Weiss 2009). This is most clearly seen in the present infinitival forms, as the different theme vowels do not always surface as such in the rest of the present paradigm. Since there are present infinitives with four different theme vowels (ignoring, at this point, some obviously irregular verbs), there are typically considered to be four different conjugational classes (typically assigned numbers 1–4):

<table>
<thead>
<tr>
<th>conjugation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>theme vowel</td>
<td>Ā</td>
<td>Ė</td>
<td>Ė</td>
<td>Ī</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONJUGATION</th>
<th>1SG.PRES.IND.ACT.</th>
<th>2SG.PRES.IND.ACT.</th>
<th>3SG.PRES.IND.ACT.</th>
<th>1PL.PRES.IND.ACT.</th>
<th>2PL.PRES.IND.ACT.</th>
<th>3PL.PRES.IND.ACT.</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFINITIVE</td>
<td>CANTĀRE</td>
<td>HABĒRE</td>
<td>SCRIBĒRE</td>
<td>DORMĪRE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1SG.PRES.IND.ACT.</td>
<td>CANTÔ</td>
<td>HABĒO</td>
<td>SCRIBÔ</td>
<td>DORMĪO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2SG.PRES.IND.ACT.</td>
<td>CANTĀS</td>
<td>HABĒS</td>
<td>SCRIBIS</td>
<td>DORMĪS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SG.PRES.IND.ACT.</td>
<td>CANTĀT</td>
<td>HABET</td>
<td>SCRIBIT</td>
<td>DORMIT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1PL.PRES.IND.ACT.</td>
<td>CANTĀMUS</td>
<td>HABĒMUS</td>
<td>SCRIBIMUS</td>
<td>DORMĪMUS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2PL.PRES.IND.ACT.</td>
<td>CANTĀTIS</td>
<td>HABĒTIS</td>
<td>SCRIBITIS</td>
<td>DORMĪTIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3PL.PRES.IND.ACT.</td>
<td>CANTĀNT</td>
<td>HABENT</td>
<td>SCRIBUNT</td>
<td>DORMIUNT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: Common (oversimplified) picture of Latin verb conjugation

This distinction gives the commonly referenced Latin “four–conjugation system”. It is generally assumed that phonological rules take care of the allomorphy in these forms, e.g. vowel length differences within the paradigm (e.g., DORMĪS ‘sleep–2SG’ vs. DORMĪT ‘sleep–3SG’) or /i/u allomorphy within the traditional 3rd conjugation present paradigm (e.g., SCRIBIT ‘write–3SG’ vs. SCRIBUNT ‘write–3PL’).

This is not to say that Latin scholars believed that there were only four ways in which Latin verbs were conjugated. There is value in the concept of “principal parts”, the set of forms that a non-native learner of a language must learn in order to be able to form all the different conjugated forms (in reality, a set of relevant conjugated forms representing different stem formations).
However, does this have any relation to how these forms are stored by speakers of Latin? People acquired some form of Latin as a first language, and there is a difference between acquiring Latin from one's surroundings as a first language and being taught it from a textbook. Regardless, this practice of classifying Latin verbs on the basis of the theme vowel (as seen in the infinitive) has been pervasive, and has definitely influenced the way that modern Romance conjugational classes are described. Moreover, this traditional description of the Latin system is by no means perfect, as is demonstrated below.

2.1.3. Introduction to issues with the four-conjugation system

Not all Latin verbs follow one of the four patterns precisely in the present system. That is, classifying verbs by their infinitive proves quite faulty after a more detailed examination of Latin data. First, not all verbs that have seemingly similar infinitival forms (i.e., verbs with the same theme vowel) are conjugated precisely in the same way, as would be required by the GDM in order for them to be considered following the same pattern.

Second, this descriptive system only attempts to account for the patterns seen in the present system of Latin verbs. There is no definitive reason to base Latin verb classification on the present stem. If we base the definition of Latin conjugational classes on the present stem, as the traditional four-conjugation system does, then we have to claim that there are several “subclasses” of each conjugation when it comes to
relating (or “linking”) the present stems to perfect stems (as well as to participial stems).

As discussed below, the perfect stem is somewhat predictable for some present classes (especially in the traditional 1st and 4th conjugations – though it is still not completely predictable), but not for others. If speakers classify verbs into classes following what the most accurate generalization would be, then why is this never considered to all be based on the perfect stem? A classification based on the perfect stem will likely yield similar (if not potentially more accurate) results in terms of predictability. That is, there may be a higher likelihood of accurately predicting a present stem from the perfect stem than vice-versa4. The reasons for not focusing on the perfect stem when categorizing verbs have nothing to do with the linguistic structure of Latin; it is due to tradition.

The bias towards the present system is seen in most accounts of the Latin verbal conjugation system. For example, Dressler (2003), in the framework discussed in the previous chapter, has his macroclasses representing the traditional Latin conjugations. From each of these macroclasses branch microclasses based on the different types of perfect stems that correspond with the present stem, and for some of these classes stem different microclasses based on the formation of the participial stem. For example, Dressler gives the following diagram for the traditional 1st conjugation:

4 It really does not matter on which stem one bases their classification, as there will always be exceptions to any widespread generalization. If the primary classification is done based on the different perfect stems, there will still be unpredictable cases, e.g., the fact that DOMUI corresponds to a present DOMĀ- would not be predictable based on the perfect stem, just as there is no way to know that DOMĀ- corresponds to a perfect stem DOMUI if the present stem is considered to be primary.
With this framework, Dressler attempts to account for the fact that the Latin conjugations are fragmented. However, there is no compelling reason to consider either stem as “basic”. It appears that even attempts to actually account for all the details (as Dressler’s does) are still biased by the traditional four-conjugation system.

Since having a diagram such as the one above implies the importance of one stem over the others in categorization, the approach that I take here diverges significantly. Instead of proposing that the present stem is dominant, and from each present stem class branch perfect stem classes (from which often branch participial stem classes), I propose that there are “linkages” between present, perfect and participial stems. In Latin, the number of present stem patterns is not immense. In addition, the number of perfect stem patterns or participial stem patterns are not numerous. However, when these linkages are taken into account, the number of possible patterns that Latin verbs may exhibit in their entire conjugation system (this includes present, perfect and participial formations) rises dramatically. These linkages are in fact patterns shown by verbs in the formation of different stems. While some of these patterns are followed by hundreds of verbs, some of the other patterns have only a few adherents. These patterns are not organized into classes. All of the patterns share some characteristics, and some definitely share more features than others.
Instead of trying to capture this with branching diagrams, I propose a model based on similarity on multiple levels.

Again, what follows does not constitute a complete listing of the possible conjugational patterns found in Latin. What the rest of this chapter does do is select several points of interest, and explains how they show the local nature of generalizations in Latin verb conjugation\(^5\).

2.2. A Closer Look at the Latin Present System

While the four-conjugational class system above is often proposed for Latin, it is definitely not the case that all verbs fall into one of these classes. Besides the fact that this only attempts to account for the present stem of Latin verbs, there are frequent cases in which certain present stems do not fit squarely into one of the four classes described above.

2.2.1. The “Capiō” class

The most notable issue with the four-conjugational class system is the existence of certain verbs that have an infinitive and several other forms that follow the same patterns as traditional 3\(^{rd}\) conjugation verbs, while having other forms that follow the traditional 4\(^{th}\) conjugation verbal conjugation pattern.

---

5 Another aspect to take into consideration here is that the data presented in this chapter is (mainly) from Classical Latin. Other dialects, particularly those that are not overtly prestigious and have no pressure towards standardization, often show a much more fragmented picture than the standard language (see Ch. 4 for Daco–Romanian, Ch. 5 for Catalan, Ch. 6 for Italian and Aromanian).
Table 2.4: Traditional 3rd vs. CAPIÔ class vs. traditional 4th

<table>
<thead>
<tr>
<th></th>
<th>3rd</th>
<th>3rd “-iō class”</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>INF</td>
<td>SCRIBÈRE</td>
<td>CAPIÈRE</td>
<td>DORMÈRE</td>
</tr>
<tr>
<td>1SG</td>
<td>SCRIBÔ</td>
<td>CAPIÔ</td>
<td>DORMIÔ</td>
</tr>
<tr>
<td>2SG</td>
<td>SCRIBIS</td>
<td>CAPIS</td>
<td>DORMIS</td>
</tr>
<tr>
<td>3SG</td>
<td>SCRIBIT</td>
<td>CAPIT</td>
<td>DORMIT</td>
</tr>
<tr>
<td>1PL</td>
<td>SCRIBIMUS</td>
<td>CAPIMUS</td>
<td>DORMIMUS</td>
</tr>
<tr>
<td>2PL</td>
<td>SCRIBITIS</td>
<td>CAPITIS</td>
<td>DORMITIS</td>
</tr>
<tr>
<td>3PL</td>
<td>SCRIBUNT</td>
<td>CAPIUNT</td>
<td>DORMIUNT</td>
</tr>
</tbody>
</table>

Table 2.4: Traditional 3rd vs. CAPIÔ class vs. traditional 4th

These verbs have been traditionally called 3rd conjugation -iō verbs' (or the “CAPIÔ class”) noting the fact the infinitives have the theme vowel Ė, like 3rd conjugation verbs (e.g., CAPÈRE, SCRIBÈRE), and a 1SG form in -iō (unlike 3rd conjugation 1SG forms, which just have -ō, e.g., CAPIÔ vs. SCRIBÔ). In sources aimed at describing the Latin system synchronically, this subgroup is usually considered part of the 3rd conjugation (Hale & Buck 1903), while in sources aimed at describing the evolution of these forms, they are usually described along with the 4th conjugation (as they have more in common historically with this class, Sihler 1995). Under the GDM, the CAPIÔ class must be considered as following a separate pattern, as the details of conjugation do not allow it to be subsumed under any other class. More specifically, the CAPIÔ class is best described as a group of patterns that all share the same pattern in the present stem, though they show variation in terms of their linkages to perfect and participial stems.

In terms of the theme vowel, this class is traditionally posited as having a theme vowel Ė, which given the normal phonological processes of Latin, gives the expected forms throughout the paradigm. However, positing a different theme vowel is tantamount to positing that this is a completely different class, no matter how superficially similar it looks to other conjugations on the surface.

Taking these verbs into account, a more accurate proposal could list five different Latin conjugational classes. While a five-conjugational class system is advantageous over the four-conjugational class system, given that the patterns
exhibited by verbs like CAPIO would be accounted for, the fact remains whose conjugation in the present system falls outside of this range of patterns.

2.2.2. Traditionally described irregular verbs

There are a number of Latin verbs whose present system conjugation diverges in some manner from any of the traditionally described conjugational classes discussed above. A few verbs appear to almost fit into one of these classes, but instead, fall a bit short. As discussed in Chapter 1, the verb DARE ‘give’ (very) superficially looks like a 1st conjugation verb. However, long ā is not found in many of the forms in which it would be expected if this were a 1st conjugation verb (e.g., INF DARE, not *dāre; 1PL DAMUS, not *dāmus, etc.). Again, an abstract underlying representation containing ā could be posited to attempt to make this pattern conform to the pattern followed by rest of the 1st conjugation present forms. However, doing so would also require the positing of a lexically-specific (i.e., non-phonological) rule that shortens this underlying /ā/ in these specific forms (e.g., /dāre/ for the infinitive). As this “rule” would be restricted to this and only this verb, doing this is in no way advantageous than just proposing that this verb just follows a unique pattern in present conjugation, which, while similar to that followed by other verbs, does show a difference that cannot be ignored. This verb has a unique present conjugation, and this uniqueness has to be somehow reflected in the grammar.

There are also patterns that diverge to a greater extent. The verb VELLE ‘to wish’ (as well as its derivatives) shows divergence from the rest of the conjugational system due to the fact it descends from a Proto–Indo–European athematic verb6. While some of

---

6 IE athematic verbs had the endings directly attached the root, e.g., *wel–mi, as opposed to ‘thematic’ verbs, which has a thematic vowel from *e/o intervening between the root and the inflectional endings.
the forms in the present paradigm were ‘thematized’\(^7\), e.g., \textsc{1sg volō} from an earlier PIE *wel-\text{-}mi (athematic), some of the other forms in the paradigm are directly descended from athematic forms, e.g., \textsc{3sg vult}, \textsc{2pl vultis}. The verb \textsc{esse} ‘to be’ shows athematic forms as well, e.g., \textsc{1sg sum} < *H\text{esmi} (although the evolution of this form, as well as that of the rest of the paradigm, is quite more complicated than it seems, see Joseph & Wallace 1987). \textsc{posse} ‘to be able’ patterns with \textsc{esse} in the present system as it is considered to be a derivative of \textsc{esse} (\textsc{posse} < *pot\text{-}esse). However, this limited generalization in the present system across these two verbs breaks down in the perfect, e.g. the \textsc{1sg.perf.ind.act.} of \textsc{sum} is fuī, while that of \textsc{possum} is potūī, not something like *potfuī (or *poffuī given the expected assimilation).

There are clearly more than four (or five!) patterns, and attempts to boil them down to four (or five) are generalizing at the expense of the actual facts. Again, as in Sanskrit, there is much more to verb conjugation in Latin than just the present systems. The Latin perfect stem also shows quite a deal of complexity, particularly when the linkages between present and perfect stems are considered.

2.3. THE LATIN PERFECT SYSTEM

At first glance, the Latin perfect system seems rather “regular”, as the perfect endings are consistent across verbs.

\(^7\) See Chapter 1 for a discussion of the selected “thematization” of this verb.
Table 2.5: Perfect indicative across traditional (present) conjugational class boundaries

<table>
<thead>
<tr>
<th>Pres.conj</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>irreg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>perf stem</td>
<td>CANTĀVĪ</td>
<td>HABŪĪ</td>
<td>SCRĪPSĪ</td>
<td>DORMĪVĪ</td>
<td>FŪĪ</td>
</tr>
<tr>
<td>1SG</td>
<td>CANTĀVĪ</td>
<td>HABŪĪ</td>
<td>SCRĪPSĪ</td>
<td>DORMĪVĪ</td>
<td>FŪĪ</td>
</tr>
<tr>
<td>2SG</td>
<td>CANTĀVISTĪ</td>
<td>HABŪIŚTĪ</td>
<td>SCRĪPSISTĪ</td>
<td>DORMĪVISTĪ</td>
<td>FUISTĪ</td>
</tr>
<tr>
<td>3SG</td>
<td>CANTĀVIT</td>
<td>HABUIT</td>
<td>SCRĪPSIT</td>
<td>DORMĪVIT</td>
<td>FUIT</td>
</tr>
<tr>
<td>1PL</td>
<td>CANTĀVIMUS</td>
<td>HABUIMUS</td>
<td>SCRĪPSIMUS</td>
<td>DORMĪVIMUS</td>
<td>FUIMUS</td>
</tr>
<tr>
<td>2PL</td>
<td>CANTĀVISTIS</td>
<td>HABUISTIS</td>
<td>SCRĪPSISTIS</td>
<td>DORMĪVISTIS</td>
<td>FUISTIS</td>
</tr>
<tr>
<td>3PL</td>
<td>CANTĀVĒRUNT</td>
<td>HABUĒRUNT</td>
<td>SCRĪPĒRUNT</td>
<td>DORMĒVĒRUNT</td>
<td>FUĒRUNT</td>
</tr>
</tbody>
</table>

While the perfect endings are consistent (even in ‘be’ Latin’s prototypical “irregular verb”), the complexity in the perfect system is located in the formation of the perfect stem itself. While the Latin system does not show as much variation as the Sanskrit system, there is still a considerable variety in the perfect formations. In addition, these different perfect formations occurs across traditional conjugational class boundaries (based on the present stem). Though there are observable tendencies (which vary in strength depending on conjugation type), the variation in perfect stem formation means that it is not possible to accurately predict the perfect stem from the present stem of a given verb. Again, as we are considering any difference (no matter how small) that is not attributable to the normal phonological rules present in the Latin to signify a different conjugational pattern, the different perfect formations that coexist with the already not-so-simple present system, leads the number of observable patterns in Latin verb conjugation to skyrocket.

2.3.1. -vī/-ūī perfects

The predominant pattern in verbs with a present stem following the traditional 1\textsuperscript{st}, 2\textsuperscript{nd} and 4\textsuperscript{th} conjugation patterns is that their corresponding a perfect stem is formed via the suffixation -vī or -ūī\textsuperscript{8}, though what it actually suffixes to varies. This is assumed to be

\textsuperscript{8} This perfect type is often considered not to have direct cognates in other IE (or Italic) languages. Some have proposed a connection between this morpheme and the ending -au seen in selected
phonologically determined depending on the sound preceding, with the allomorph \(-\text{v}i\), following a vowel, e.g., \(\text{AMAVI} '\text{love 1SG.PERF.IND.ACT}'\), and the allomorph \(-\text{u}i\) following a consonant, e.g., \(\text{HABUI} '\text{have 1SG.PERF.IND.ACT}'\). While this allomorphy is phonologically predictable, the presence of this triggering vowel is not. For example, most traditionally described 1\textsuperscript{st} conjugation verbs have a perfect in \(-\text{Av}i\) (the theme vowel \(\text{A}\) is seen here, as it is in the present system). However, there is a small number of verbs that form the perfect by adding the same suffix directly to the root, and thus it shows up as \(-\text{u}i\) as it follows a consonant, rather than a vowel, e.g., the perfect of \(\text{DOMO} '\text{tame}'\) is \(\text{DOMUI}\), rather than \("\text{domavi}, cf. INF \text{DOMARE}\). Thus, while the perfect stem of 1\textsuperscript{st} conjugation verbs is usually predictable from the present stem, this is definitely not always the case. This can again be explained in diachronic terms, as this smaller class of 1\textsuperscript{st} conjugation verbs has a different origin (usually inherited verbs from IE) than the more frequent types (usually denominatives, iteratives, etc., DeWandel 1982). However, this distinction is only discovered through comparative Indo-European linguistics, something of which no Latin speaker would have been aware.

So, given only this distinction, the traditionally-described 1\textsuperscript{st} conjugation must be treated as two separate patterns, as even though verbs show present systems following the same pattern, there are differences in their respective perfect systems.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>PRES. stem in</th>
<th>3SG.PRES.IND.ACT</th>
<th>PERF. stem in</th>
<th>3SG.PERF.IND.ACT.</th>
<th>'sing'</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(-\text{A})</td>
<td>CANTAT</td>
<td>(-\text{Avi})</td>
<td>CANTAVIT</td>
<td>'sing'</td>
</tr>
<tr>
<td>B</td>
<td>(-\text{A})</td>
<td>DOMAT</td>
<td>(-\text{u}i)</td>
<td>DOMUIIT</td>
<td>'tame'</td>
</tr>
</tbody>
</table>

Table 2.6: Traditional Latin 1\textsuperscript{st} conjugation verbs, different PERF patterns

This actually is not the extent of the patterns covered by the traditional concept of the Latin 1\textsuperscript{st} conjugation. Some of the perfect stem formation patterns that more frequently perfect endings in Vedic (Weiss 2009), as well as with similar forms in Anatolian and Tocharian (Markey 1979)

---

62
correspond to other present tense classes (see the following sections) also are found to a much more restricted extent linked with present stems in -ā. For example, stō ‘stand’ (3SG.PRES. STAT) has a reduplicated perfect (though the reduplication has been obscured, e.g., 1SG.PERF. STETI), while iuvō ‘help’ (e.g., 3SG.PRES. IUVAT) has a long vowel perfect (e.g., 1SG.PERF. IŪVI). Even if all 1st conjugation verbs follow the same pattern in the present system, there are at least four different ways of forming the perfect. So, taking the GDM seriously, these must be considered four different patterns) once one takes into account the present stem – perfect stem linkages (and this does not even consider the participial stem).

This is a trend that is seen over and over again in Latin verbs. Just because a group of verbs show a present stem following a given pattern, it does not necessarily mean that all corresponding perfect stems of this group of verbs will all be formed following the same pattern. While the 1st conjugation seems somewhat “unified” in the present system, when comparing it to the perfect stem, it begins to look somewhat more fragmented (in the sense of Joseph & Janda 1986). When considering these linkages. The fragmentation of the other conjugational classes once these linkages are taken into account dwarfs that seen in the 1st conjugation.

2.3.2. Other perfect stem types

While the perfect stem patterns in vī /–uī are shared by the majority of Latin verbs, there is a variety of other patterns seen predominantly linked to traditional 3rd conjugation verbs, but also elsewhere to a more limited extent. Unlike the pattern in – vī /–uī, these perfect formations are found commonly in other Indo-European languages in aorist/perfect formations (the Latin perfect can be seen as a ‘mixture’ of the PIE perfect and aorist, see Sihler (1995)). Besides the common type seen above, the following types of perfects are also attested in Latin:
While these perfect formation patterns are found predominantly linked to traditional 3rd conjugation present stems, they are also found to a more limited extent linked to other present stem patterns.

<table>
<thead>
<tr>
<th>1SG.PRES</th>
<th>1SG.PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>IUVÔ</td>
<td>iÜVÔ</td>
</tr>
<tr>
<td>STÔ</td>
<td>STETÎ</td>
</tr>
<tr>
<td>CAVÉÔ</td>
<td>CĂVÎ</td>
</tr>
<tr>
<td>MORDEÔ</td>
<td>MOMORDÎ</td>
</tr>
<tr>
<td>AUGEÔ</td>
<td>AUXÎ</td>
</tr>
<tr>
<td>VINCIÔ</td>
<td>VÎNXÎ</td>
</tr>
<tr>
<td>VENIÔ</td>
<td>VÊNI</td>
</tr>
</tbody>
</table>

Table 2.8: Non 3rd conjugation with alternative perfect stem formation

While the vast majority of Latin verbs follow one of these patterns, there are several verbs with perfect stem formations that show patterns that cannot really be subsumed under any of the categories described above. For example, some verbs have suppletive perfect stems, e.g., 1SG.PRES SUM ~ 1SG.PERF FUI ‘be’ and 1SG.PRES FERÔ ~ 1SG.PERF TULÎ ‘carry’. While these patterns can often be satisfactorily explained diachronically (some verbs just did not have a perfect tense formation in PIE, including the ancestor of FERÔ, Sihler 1995), speakers of Latin did not have access to this historical and comparative information.

An observation of the linkages the different present stems with the different options for perfect tense formation shows that the Latin conjugational class system is
considerably more complicated than superficially on the surface. Again, this is only considering the present and perfect stems; additionally taking the participial system into account shows an even greater level of fragmentation.

2.4. **The Participial Stem**

The Latin past passive participle is generally formed via the suffixation of -TO. In many verbs, it appears as if the suffix is added directly to the present stem (e.g. PRES. CANTĀ-, PART CANTĀTU- ‘sing’; PRES. DORMĪ-, PART DORMĪTU- ‘sleep’), while for others, this is not the case. For this reason, the presence of a third stem type has to be accounted for (the same reason why learners of Latin today have to memorize the 4th principal part).

While it is definitely not possible to say that the material from which the participial stem is built is equivalent to another stem, there is a somewhat strong connection (or as Hale and Buck 1903 put it, “a certain relationship”) between the participle and the perfect, which holds in most, but not all situations:

<table>
<thead>
<tr>
<th>perfect type</th>
<th>participle type</th>
<th>ISG.PERF</th>
<th>PPP</th>
<th>‘word’</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfects in –AVI</td>
<td>participle in –ĀTUS</td>
<td>AMĀVĪ</td>
<td>AMĀTUS</td>
<td>‘love’</td>
</tr>
<tr>
<td>perfects in –IVI</td>
<td>participle in –ITUS</td>
<td>AUDĪVĪ</td>
<td>AUDĪTUS</td>
<td>‘hear’</td>
</tr>
<tr>
<td>perfects in –UL</td>
<td>participle in –ITUS</td>
<td>DOMŪI</td>
<td>DOMITUS</td>
<td>‘tame’</td>
</tr>
<tr>
<td>other perfects (long vowel, redup, etc.)</td>
<td>participle in –SUS (following dentals)</td>
<td>MOMORDĪ</td>
<td>MORSUS</td>
<td>‘bite’</td>
</tr>
<tr>
<td></td>
<td>participle in –TUS (elsewhere)</td>
<td>FĒCĪ</td>
<td>FACTUS</td>
<td>‘make’</td>
</tr>
</tbody>
</table>

Table 2.9: Generalizations of correspondences between perfect and participial stems

While these generalizations cover most of the cases, there exceptions to each. There are some verbs that have perfects in –AVI that have past participles in –TUS, e.g., PERF PŌTĀVĪ – PART POTUS, ‘drink’ (though an alternate form expected by the generalizations above, PŌTĀTUS, also exists). There are also a few verbs that have perfects in –UL but a
participle in –TUS (e.g., DOCUĪ – DOCTUS) rather than a form in –ITUS as expected by the generalizations above.

In addition, with non –vī/-ūī perfects, the conditions under which a given allomorph (–TUS or –SUS) appears on the surface are not completely consistent as there are some forms that have –SUS following a non-dental, e.g., following a labial in LĀP–SUS ‘slip’ or following a velar in FIXUS (fik–sus) ‘fix’. These are undoubtedly the result of analogy to other participles in –SUS (or potentially analogy to a perfect stem in –Sī). Regardless, at the moment at which –SUS was extended to a root without a final dental consonant, this rule, which was earlier apparently completely phonetically motivated, was no longer predictable. Since the distribution of these allomorphs is no longer predictable, it warrants the classification of these verbs into separate classes under the view used here. Analogy is often (and somewhat erroneously) considered to be a simplifying force. However, in this case, analogy took something cohesive and phonetically motivated and fragmented it completely. This is similar to the cases of German Umlaut and Sanskrit Reduplication (Joseph & Janda 1986) discussed in the previous chapter.

In addition, there are some cases where the participle shows a long vowel in the root. All of the attested cases of verbs having long vowel participles also have long-vowel perfects (e.g. PERF RĒXI ~ PART RĒCTUS). However, it is not the case that all long vowel perfects have a corresponding long vowel in the participle (e.g. PERF RŪPI ~ PART RUPTUS). Thus, again, as this is not completely predictable, these must be considered different patterns under this approach.

2.5. Other Patterns

Some verbs that are traditionally described as being part of the 3rd conjugation that
show variant patterns involved in the formation of the present stem from the root. The
treatment in Hale & Buck (1903) is notable in that they breaks down their vision of the
3rd conjugation in terms of the different present stem types observed.

<table>
<thead>
<tr>
<th>Present stem type</th>
<th>e.g.,</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Simple thematic present</td>
<td>SCRIBÒ 'write'</td>
</tr>
<tr>
<td>B Present with reduplication</td>
<td>SERÔ9 'sow'</td>
</tr>
<tr>
<td>C Present with infixed nasal</td>
<td>TANGÔ 'touch'</td>
</tr>
<tr>
<td>D Present in –NÔ</td>
<td>STERNÔ 'spread'</td>
</tr>
<tr>
<td>E Present in –TÔ</td>
<td>FLECTÔ 'bend'</td>
</tr>
<tr>
<td>F Present in –SCÔ</td>
<td>PASCÔ 'feed'</td>
</tr>
<tr>
<td>G Present in –ESSÔ</td>
<td>CAPESSÔ 'seize'</td>
</tr>
<tr>
<td>H Present in –UÔ</td>
<td>STRUÔ 'pile'</td>
</tr>
<tr>
<td>I Present in –IÔ10</td>
<td>CAPIÔ 'take'</td>
</tr>
</tbody>
</table>

Table 2.10: Hale & Buck’s (1903) breakdown of the traditional Latin 3rd conjugation

These “subclasses” of the third conjugation show completely “regular” present tense
formations, i.e., they conjugate regularly with no unexpected stem allomorphy within
the present system. The issue here is that, in many of these different present stem
formations, there is material present in the formation of the present stem (e.g., an
infixed nasal in Hale & Buck’s class C, etc.) whose appearance is not consistent across
the perfect and participial stems. Thus, this is another aspect of Latin verbal
conjugation that needs to be recognized when discussing the different patterns. Several
issues brought up with discrepancies between present/perfect/participial stem material
are discussed below.

9 From PIE *si–sH₁–e– (Weiss 2009). This, and other examples of the reduplicated present show
that these were probably not derived synchronically from the root; there must be some sort of rule
that links this present stem with the appropriate perfect. Latin speakers were likely not doing any
synchronic reduplicating in the present system; this probably was just another present stem that
was linked to a perfect stem that looked somewhat similar as some of the same material exists.
10 This is the “CAPIÔ class” discussed above, which Hale & Buck consider to be part of a rather
broad 3rd conjugation.
2.5.1. Present stems with an “infixed” nasal consonant

Among of the types of traditionally described 3\textsuperscript{rd} conjugation verbs is one that forms the present stem through the infixation of a nasal consonant before the final consonant of the root (e.g., Lat FING- ‘form’ < PIE *dheiǵh-, cf. Grk. τείχoς). The appearance of this nasal consonant, while consistent throughout the present system (e.g., FINGŌ ‘1SG.PRES.SUBJ.ACT’, FINGOR ‘1SG.PRES.SUBJ.PASS’, FINGĒBAM ‘1SG.IMP.IND.ACT’, FINGENDUM ‘GER’, etc.), does not show the same distribution in the perfect and participial systems. Strictly etymologically, the nasal should only occur in the present system, and in some verbs, this is indeed what is found. In other verbs, however, the nasal has been extended from the present system to the perfect system, but not to the participle. There are also cases where the opposite is true; the nasal has been extended to the participle, but not to the perfect. And finally, there are some cases in which the nasal has been extended to all forms of the verb (from the present system into the perfect and the participle).

<table>
<thead>
<tr>
<th></th>
<th>present system (1SG.PRES.IND.ACT)</th>
<th>perfect system (1SG.PERF.IND.ACT)</th>
<th>past participle</th>
</tr>
</thead>
<tbody>
<tr>
<td>nasal only in present stem</td>
<td>TANGŌ</td>
<td>TETĪGĪ</td>
<td>TACTUS</td>
</tr>
<tr>
<td>nasal extended to perfect, but not to participle</td>
<td>FINGŌ</td>
<td>FINXĪ</td>
<td>FICTUS</td>
</tr>
<tr>
<td>nasal extended to participle, but not to perfect</td>
<td>PUNGŌ</td>
<td>PUPUGĪ</td>
<td>PUNCTUS</td>
</tr>
<tr>
<td>nasal extended to all forms</td>
<td>IUNGŌ</td>
<td>IŪNXĪ</td>
<td>IŪNTUS</td>
</tr>
</tbody>
</table>

Table 2.11: Nasal infix distribution (shaded cells indicate presence of the nasal)

Under the definition of conjugational pattern presented here, the first three verbs (TANGŌ, FINGŌ and PUNGŌ) in the table above must be considered to follow distinct conjugational patterns, as there are differences across the present, perfect and
participial stems as to the appearance of the nasal. The fourth example (IūNGŌ), while showing a different distribution of the nasal from the previous three, could just be considered a non–nasal verb, as there is no alternation as to its appearance. We know that it is a diachronically a nasal–infix verb (cf. Skt. present stem yu–na–k–, aorist stem ayuj–), but, to speakers of Latin, it was probably just considered another third conjugation verb with an s–perfect. Latin did have the noun IUGUM 'yoke', and one might propose that this noun and the corresponding verb are both derived synchronically from the same root, but saying this does not gain anything and would lead to the necessity of proposing a lexically specific nasal infix rule that would apply to all the inflected forms of the related verb IUNCO.

Note that while there is variation as to the appearance of the nasal in these forms (whether it has been extended to the perfect system and past participle or not), there are other verbs descended from this same group that do not show such variation. There is group of verbs with a 1SG present in –NŌ, which is actually a nasal infix present stem that was originally laryngeal–final (e.g., PIE *sperH₁– → *sper–n–H₁ → Lat. SPERN–, Sihler 1995); these apparently never have the nasal extended to the perfect or participle.

<table>
<thead>
<tr>
<th>1SG.PRES.IND.ACT</th>
<th>1SG.PERF.IND.ACT</th>
<th>PPP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SPERNŌ</td>
<td>SPRĒVĪ</td>
<td>SPΡΤUM</td>
<td>'strew'</td>
</tr>
<tr>
<td>PELLO₁¹¹</td>
<td>PEPLĪ</td>
<td>PULSUM</td>
<td>'drive'</td>
</tr>
</tbody>
</table>

Table 2.12: Lat. verbs in –NŌ (pres. stem from PIE *CV–n–H)

2.5.2. Presents in –scō

Another set (broadly defined) of verbs that show some sort of stem extension is that with a present system showing the suffix –scō. The Latin affix –ESC/-ISC¹² gives

---

¹¹ The result of assimilation from *pel–nō
¹² This affix comes from PIE *sḱ`/₁ (difficult to reconstruct semantically, as its reflexes throughout the Indo–European languages have a variety of functions (e.g. present stem formant in Sanskrit, non–active marker in Albanian, etc.)
inchoative or inceptive meaning when attached to verbs (e.g., DORMĪCŌ ‘fall asleep’
cf. DORMĪ ‘sleep’, or denominative/deadjectival function when attached to
noun/adjectives (e.g., CLĂRĒSCŌ ‘become bright’, cf. CLĂRUS ‘bright’, Meul 2010). While
this affix consistently appears in the present system of these verbs, it, for the most part,
does not appear in the perfect system, and the corresponding perfect stem would be the
same as the verb from which it is derived)

<table>
<thead>
<tr>
<th>1SG.PRES.IND.ACT</th>
<th>1SG.PERF.IND.ACT</th>
<th>'sleep'</th>
</tr>
</thead>
<tbody>
<tr>
<td>DORMĪO</td>
<td>DORMĪVĪ</td>
<td></td>
</tr>
<tr>
<td>DORMĪCŌ</td>
<td>DORMĪVĪ</td>
<td>'fall asleep'</td>
</tr>
</tbody>
</table>

*Table 2.13: Present/Perfect of inchoative verbs*

In the case of verbs in –SCŌ that are formed from nouns or adjectives, their
corresponding perfect system is normally “follows the type which is commonest in those
derived from verbs”, for example, –ĀVĪ for presents in ĀSCŌ, e.g., VESPERĀSCŌ ‘become
evening–1SG.PRES.’ (<VESPER ‘evening’), VESPERĀVĪ ‘1SG.PERF’, etc. (Hale & Buck 1903).

In addition to these derived verbs, there is a small group that was inherited into
Latin with the affix already attached and do not necessarily have inchoative or inceptive
meaning. Despite the fact that they were inherited into Latin with the suffix already
attached, the suffix normally only appears in the present system, and is absent in the
perfect.

PRES NŌSCŌ – PERF NŌVĪ ‘know’

cf. Alb. njoh ‘know–1SG’ (where –h < PIE *–sk)

However, in the case of POSCŌ ‘ask’ the affix has been extended from the present stem
to the perfect and participial stems.
Thus, while these two verbs have similar origins (in terms of the presence of the suffix), they must be considered to be following different patterns synchronically in Latin. As the suffix has been extended across-the-board, POSCÔ can just be considered another e-stem with a reduplicated perfect. Alternatively, NOSCÔ and the derived verbs must be considered to follow a distinct pattern (or, more accurately, in a variety of patterns, given the differences in linkages with perfect formations) due to the suffix’s absence outside of the present system.

The above two sections are only two examples where there is a discrepancy between the material in the present and perfect stems. This is not the extent of these differences, as other present stem types show similar patterns. This just goes to show that there really is a wide variety of patterns in Latin conjugation, something that is visible especially when one compares the present system with that of the perfect.

2.5.3. Deponents & Defectives

One final point of Latin verb conjugation that must be considered pertains to the groups of Latin verbs that are deponent or defective.

Defective verbs are those that are missing certain forms. For example, there is a small set of Latin verbs that generally only have morphologically perfect forms. Even though these forms are morphologically perfect, they are semantically present:

| 1SG.PRES. | ÓDÌ      | ‘I hate’ |
| 2SG.PRES. | ÓDISTÌ   | ‘you hate’ |
| 3SG.PRES. | ÓDIT     | ‘(s)he hates’ |

Table 2.14: Defective verb (no morphological present tense)
Even though these are morphologically perfect forms, they are used in present contexts. As their conjugation follows a distinct pattern, verbs such as these have to be considered a belonging to distinct patterns.

Latin deponent verbs are morphologically passive in the present system but semantically active. This particularity is attested in other Indo-European languages as well. For example, the following is the present indicative ACTIVE of the verbs MĪROR 'admire' and PARTIOR 'share'. Deponent verbs, for the most part, are regular. However, again, they are semantically active and morphologically passive. Again, this means we have another pattern (really, several new patterns, as there are several types of deponent verbs, e.g., the differences in conjugation between MĪROR and PARTIOR).

<table>
<thead>
<tr>
<th></th>
<th>morphologically passive semantically active (DEPONENT)</th>
<th>morphologically passive semantically passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>MĪROR</td>
<td>PARTIOR</td>
</tr>
<tr>
<td>2SG</td>
<td>MĪRĀRIS</td>
<td>PARTĪRIS</td>
</tr>
<tr>
<td>3SG</td>
<td>MĪRĀTUR</td>
<td>PARTĪTUR</td>
</tr>
<tr>
<td>1PL</td>
<td>MĪRĀMUR</td>
<td>PARTĪMUR</td>
</tr>
<tr>
<td>2PL</td>
<td>MĪRĀMINĪ</td>
<td>PARTĪMINĪ</td>
</tr>
<tr>
<td>3PL</td>
<td>MĪRANTUR</td>
<td>PARTIUNTUR</td>
</tr>
</tbody>
</table>

Table 2.15: Latin deponent verbs vs. PASS.PRES conjugation

The perfect system of deponent verbs is formed with the past passive participle plus a conjugated form of SUM. Again, this is precisely how the passive perfect is formed with non-deponent verbs, e.g., CANTĀTUS SUM ‘sing–1SG.PERF.IND.PASS’, MĪRATUS SUM ‘admire–1SG.PERF.IND.ACT’. In addition, there are some verbs that have a morphologically and semantically active present stem that have a morphologically passive but semantically active passive system (often referred to as “semi-deponent verbs”, e.g., FĪDĪ ‘trust–1SG.PRES.IND.ACT, FĪSUS SUM ‘trust–1SG.PERF.IND.ACT’ (again, semantically active, but
morphologically passive). Again, these verbs belong to a separate pattern as there is a distinct linkage between the present and perfect systems.

2.6. **Summary and Conclusions**

This chapter has aimed to discuss issues relating to Latin conjugation in the framework described in the introductory chapter of this dissertation. Even though the Latin conjugational system is usually described as being broken down into four or five conjugations, there are clearly many more different patterns actually shown in Latin verb conjugation.

Even if we take the traditional four-conjugation system as only attempting to discuss the Latin present system, it still comes up short, as there are verbs whose present inflection does not match any of these four conjugations (e.g., deponent verbs, the *capiō* class, *sum*, *volō*, etc.). Once one also considers the different perfect stem formations, which are not predictable from the present stem, the system gets more complicated, as there has to be a series of “linkages” between the different stems of a single verb. And, when one looks at the different patterns formed by these linkages, the observation that these generalizations are actually quite local becomes clear.

It is true that in the development of Latin into Romance, some of the complexity of the Latin verb has been lost (this same statement could be made for the evolution of Latin from Indo-European). Some ‘irregularities’ have been leveled out and some inflectional verbal categories are gone completely; e.g., the Latin passive and future have disappeared, replaced by (originally) periphrastic constructions. However, this is not to say that the Romance conjugational system is any less complicated than that of Latin. Some simplification has definitely occurred, but over time, other factors have led to increased fragmentation. Chapter 3 gives an overview of the evolution of the Romance conjugational system from Latin, and show how some of the Latin
‘irregularities’ being eliminated (e.g., the Romance trend moving towards regularization of the perfect and past participle) is leveled out by increasing complicating factors that have arisen in Romance (e.g., fragmentation of the traditional 1st conjugation in Alghero Catalan).
CHAPTER 3:
INTRODUCTION TO THE DEVELOPMENT OF THE ROMANCE VERBAL SYSTEM(S)

3.0. INTRODUCTION
The present chapter discusses general issues regarding the evolution of Latin morphology into that of the Romance languages. On the surface, it seems as though Romance verbal morphology has been considerably simplified when compared to that of its ancestor. This can especially be seen in the loss of several verbal categories, e.g., the entire Latin morphological passive system, as well as in various instances of leveling both within and between paradigms.

However, just as been discussed in the past two chapters regarding the morphological change does not necessarily progress unidirectionally towards a greater level of simplicity. Along with the various “simplifications” seen in the development of Romance, a number of “complications” can also be observed. Several of these “complications” are briefly discussed below, and two specific cases that will be discussed in much greater detail in Chapters 4 and 5 are introduced here.

3.1. LOSS OF COMPLEXITY
3.1.1. Loss of Verbal Categories
The Romance Languages are typically considered to be morphologically simpler than their ancestor. This, in some regards, is true, as some Latin morphological categories have been lost completely in the development into Romance. Within noun morphology, this is dramatically seen in the loss of all case distinctions (except in pronouns) across
Modern Romance\(^1\), with the exception of Balkan Romance, which maintains a limited amount of Latin case morphology (e.g., DR *casā* ‘house–NOM/ACC’ vs. *case* ‘house–GEN/DAT’), and Rhaeto–Romance varieties that maintain limited case distinctions in adjectives (Haiman & Benicà 1992). While not as dramatic, some morphological categories have also been lost in the development of the Romance verb. For instance, the passive conjugation has been completely lost. As discussed in Chapter 2, inflected active verb forms had a corresponding passive form where the difference was expressed morphologically via a different set of verbal endings. This morphological distinction is completely absent in Romance, where passive voice is expressed using a periphrastic construction.

The Latin future tense formation was also completely lost in the development to Romance. Latin had a synthetic future tense whose formation followed two patterns roughly dependent on traditional conjugational class\(^2\), e.g., Lat. *cántābō* ‘sing–1SG.FUT’, *vēniām* ‘come–1SG.FUT’. This formation has been completely lost, and has been replaced by a number of (at least, originally) analytic formations in Romance. The most widespread comes from a construction consisting of the infinitive plus a conjugated form of *habeō*, e.g., *cántāre* habeō  \(\rightarrow\) Ptg. *canterei*, Sp./Cat. *cantaré*, Fr. *chanterai*, It. *canterò*, etc. ‘sing–1SG.FUT’. Other formations include a conjugated form of ‘have’ plus the preposition *a* followed by an infinitive, e.g., Sard. *appu a kantarī*\(^3\), and one consisting and an invariant particle descended from the verb ‘want’ followed by the present subjunctive of the verb, e.g., AR *va s–cāntu*\(^4\), among others.

---

\(^1\) This is not to say that this loss took place at the same rate in all of the languages, e.g., Medieval French maintained case distinctions longer than most other Romance languages.

\(^2\) Futures in –ē– are found in the traditional 1\(^{\text{st}}\) and 2\(^{\text{nd}}\) conjugations. The traditional 3\(^{\text{rd}}\) and 4\(^{\text{th}}\) conjugations have stems in –ē– except for the 1SG, which has a stem in ā. However, futures in ā are attested in earlier texts in the 4\(^{\text{th}}\) conjugation as well (Hale & Buck 1903).

\(^3\) Where *appu* is the 1SG.PRES of ‘have’ and *kantarī* is the infinitive of ‘sing’.

\(^4\) The Balkan Romance future likely has its origins in contact with other languages of the Balkans, most specifically Greek, which have similar patterns, e.g., Grk. *θα πίνω*, Mk. *k’e pijam*, ‘drink–1SG.FUT’ where *θa/k’e* are from participles descended from ‘want’, and *πίνω/pijam* are the 1SG forms of ‘drink’.
3.1.2. Loss of stem allomorphy

There has been a loss of some stem allomorphy both within and across paradigms in the development of Romance from Latin. Some verbs that showed some “irregularity” (or, divergence from the more common patterns) have been ‘regularized’; e.g., as discussed in Chapter 1, the –ss–/–T– allomorphy of the present tense paradigm of Lat. POSSUM ‘be able’ is gone in Spanish and Daco–Romanian, in favor of a form with the reflex of –T–.

<table>
<thead>
<tr>
<th>Lat.</th>
<th>Sp.</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.PRES.IND. POSSUM</td>
<td>puedo</td>
<td>pot</td>
</tr>
<tr>
<td>2SG.PRES.IND. POTES</td>
<td>puedes</td>
<td>poti</td>
</tr>
<tr>
<td>3SG.PRES.IND. POTEST</td>
<td>puede</td>
<td>poate</td>
</tr>
</tbody>
</table>

Table 3.1: Leveling of present paradigm of Lat. POSSE in Sp. & DR

However, as discussed earlier, this was not a Pan–Romance development, as Portuguese and Italian maintain some of this allomorphy (e.g., Lat. POSSUM → It, Ptg. posso).

In many of the verbs that showed different stem material in the present and perfect systems discussed in Chapter 2, this stem allomorphy has often been eliminated (to varying degrees) in Romance. This can be seen in the verbs that were inherited into Latin with the –sc– affix, e.g., Lat. COGNOSCŌ ‘know–1SG.PRES.IND.ACT.’, that did not show the affix in the perfect system, e.g., Lat. COGNŌVĪ ‘know–1SG.PERF.IND.ACT.’.

However, this present–perfect allomorphy has been eliminated in some of Romance: e.g., Sp. conocí, AR conoscui, ‘know–1SG.PERF.’ etc. Also recall the Latin ‘nasal presents’, which had a nasal consonant in the present stem that had been extended variably (if at all) to the perfect and participle. In some cases, the nasal has been

5 The PRES–PERF allomorphy in this verb is maintained in Std. It., e.g., 1SG.PERF conobbi. However, the whole paradigm does not have the same root, as in Italian perfects, the 1SG/3SG/3PL normally inherit the Latin perfect, while the 2SG/1PL/2PL are regularized (they show the present stem, e.g., 1PL PERF conoscemmo). See Chapter 6 for a discussion of this distribution in Sicilian dialects.
extended in Romance to all forms (similar to what happened to IUNGÔ 'join' in Latin),
e.g., Lat. vinçô 'win–1SG.PRES.', viçî 'win–1SG.PERF', victus 'win–PPP' give Sp. vîgzo,
vençî, vencido, and It. vinco, vînsi, vînto, respectively. However, complete extension to
all forms is not always the result, as there are cases where the situation is maintained as
it was in Latin, e.g., Lat. rumpô 'break–1SG.PRES', rûpî 'break–1SG.PERF', ruptus 'break–
PPP' give It. rompo, ruppi, rotto, respectively, with no further extension of the nasal
outside of the present system. In Spanish, on the other hand, the evolution of this verb
involves the extension of the nasal to the preterite, but not the participle, e.g., Sp.
rompo, rompî, roto.

Leveling across paradigms also occurred in the development of Romance (again,
if we consider the present and the perfect to be two paradigms that are somewhat
connected by a sort of linkage rule). There is evidence for this in that many verbs
changed their perfect formation throughout Latin, as well as into Romance; for example,
Latin reduplicated perfects are long lost in Italian (e.g., Lat. momordî → It. morsî6 'bite–
1SG.PERF.IND.'). However, not all of the differences in perfect formation are lost from
Latin, i.e., not everything is leveled to a single patterns. Some of these forms have the
expected phonological development from Latin, e.g., Lat. dîxi → It. dissi 'say–
1SG.PERF.IND.,'

Since the morphological passive was eliminated in the evolution of Romance,
deponent verbs (those which have a present tense that is semantically active but
morphologically passive) have evolved to follow the pattern of non–deponent verbs, as
seen in the Spanish and Daco–Romanian outcomes of the Latin deponent verb morior
'die'.

---

6 The presence of an /s/ in the Italian form is another story. Several Latin verbs that had
reduplicated or long vowel perfects underwent analogy in the development to Italian.
Interestingly, these verbs often do not end up in Italian using the default, normal perfect (from
Latin -vi), but rather an s–perfect, e.g., Lat. momordî → It. morsî; Lat. viçî → It. vînsi. This is not
always the case, as some long vowel perfects have other outcomes.
The same trend is true for Latin semi-deponent verbs (those with a typical present system but a perfect system that was morphologically passive/semantically active perfect). Again, since the passive conjugation was lost, semi-deponents switched to follow the predominant pattern, as shown in the development of the perfect of Lat. GAUDEÔ ‘rejoice’ (cf. 1SG.PERF.IND.ACT GÂVISUS SUM) into Italian.

There are various other examples of morphological ‘simplification’ in the evolution of Romance verbal system. However, what specifically concerns us here are potential trends of simplification in the conjugational class systems from Latin to the various Romance varieties.

3.1.3. Development of Latin “conjugational classes”

Just as most descriptions of Latin conjugational classes are based squarely on the theme vowel (distinctions in which are seen most clearly in the infinitive), most descriptions of Romance languages follow the same general principle. Verb classes, according the most grammars of Romance, are determined by the theme (or infinitival) vowel (e.g., Badia i
Margarit 1994 for Catalan, Pei 1941 for Italian, Capidan 1925, 1932 for Balkan Romance, etc.). In Latin, this gives the traditional four-conjugation system, which, as discussed earlier, is quite flawed when the specific details are examined. While the conjugation to which Romance verbs pertain is also traditionally defined by their theme vowel, there are usually fewer than four conjugational classes proposed due to sound changes that occurred in the development of Romance. Most notably, the loss of vowel length distinctions collapsed the infinitival endings of the former traditional Latin 2nd (verbs in –ERE) and 3rd (verbs in –ERE) conjugations. This is not to say that these two conjugations are always “collapsed”, as the stress difference in infinitives of the 2nd and 3rd conjugations was maintained in Italian, Catalan and elsewhere (e.g., Lat. TENERE → It. tenere, Cat. tenir; Lat. SCRIBERE → It. scrivere, Cat. escriure); some finite forms showing this stress distinction are also seen in most varieties Balkan Romance (e.g., DR putēm/putēti ‘can–1PL/2PL.PRES.IND’ vs. mérgem/mérgeti ‘go–1PL/2PL.PRES.IND’).

The following chart shows the evolution of the traditional Latin four-conjugalional class system into the various Romance languages. Note that the following chart does not imply in any way that the inventories of these classes are maintained from Latin to Romance, as there have been frequent shifts between conjugational patterns. Rather, all it serves to do is show how the flawed traditional Latin four conjugational class system developed into the flawed traditional Romance two/three/four conjugational class systems.

---

7 The general stress pattern in Latin for trisyllabic words was that if the penultimate syllable was heavy, it was stressed. If the penultimate syllable was light, then the antipenultimate syllable was stressed. Since 1st, 2nd and 4th conjugation infinitives had an ending in a long vowel (–ARE, –ERE, –IRE, respectively), they have penultimate stress. 3rd conjugation infinitives, on the other hand, had an ending in a short vowel (–ERE), thus a light penultimate syllable, and thus, had antipenultimate stress.
Table 3.4: Evolution of Lat. infinitival endings in selected Romance languages

<table>
<thead>
<tr>
<th></th>
<th>-ARE</th>
<th>-ERE</th>
<th>-ERE</th>
<th>-IRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ptg.</td>
<td>-ar</td>
<td>-er</td>
<td>-ir</td>
<td></td>
</tr>
<tr>
<td>Sp.</td>
<td>-ar</td>
<td>-er</td>
<td>-ir</td>
<td></td>
</tr>
<tr>
<td>Cat.</td>
<td>-ar</td>
<td>-er/-re/-r</td>
<td>-ir</td>
<td></td>
</tr>
<tr>
<td>Fr.</td>
<td>-er</td>
<td>-re</td>
<td>-ir</td>
<td></td>
</tr>
<tr>
<td>It.</td>
<td>-are</td>
<td>-ere</td>
<td>-ire</td>
<td></td>
</tr>
<tr>
<td>Sic.</td>
<td>-ari</td>
<td>-iri</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DR</td>
<td>-a</td>
<td>-ea</td>
<td>-e</td>
<td>-i</td>
</tr>
</tbody>
</table>

In most of Romance, due to the collapse of Ė and Ė, what were the traditional Latin 2\textsuperscript{nd} and 3\textsuperscript{rd} conjugations are considered to have collapsed into a single conjugation. For some of the languages in which this collapse occurred, there are proposals (but they are less frequent) that maintain four conjugational class systems. For example, some accounts of Italian verb morphology propose four classes because, as discussed above, the infinitives from Latin –ERE and –ERE maintain a distinction, as those from –ERE maintain stress on the ending, while those in –ERE maintain stress on the root. However, the most common analysis is to group all Italian verbs in –ere into a single conjugation, regardless if stress falls on the ending or the root.

This “collapse” looks, rather superficially, more extreme in some Sicilian and Sardinian dialects of Italian\textsuperscript{8}, for which some sources claim only two conjugations, the first being the descendant of the Latin 1\textsuperscript{st} conjugation, and the second covering the rest of the traditionally described Latin conjugational system. However, this “collapse” is solely based on the infinitive of these verbs, the forms of which (in particular, the –ERE, –ERE and –IRE infinitival endings) have been collapsed as the result of regular sound changes\textsuperscript{9}. In all of these cases, however, there are far more than two patterns. One

\textsuperscript{8} By “Sardinian dialects of Italian” I refer to the dialects of northern Sardinia, which along with the dialects of Corsica, are normally considered very divergent Tuscan dialects of Italian (Tagliavini 1963). There is also a “Sardinian language” (definitely not a dialect of Italian) spoken in the central and southern parts of the island.

\textsuperscript{9} Latin Ė and Ė give Sicilian i regardless of context, and Latin Ė gives Sicilian i in unstressed position (the stress distinction in infinitival forms is maintained), thus giving the ending –iri for the descendants of the Latin 2\textsuperscript{nd}/3\textsuperscript{rd}/4\textsuperscript{th} conjugations (Tagliavini 1963).
example of this can be seen in Bazzoni’s (1999) description of the Sassarese dialect of Italian. He describes the Sassarese conjugational class system as follows:

In Sassarese, different from Italian, we only have two conjugations. To the first pertain verbs with infinitives in –a, corresponding to Italian –are. To the second pertain verbs with the infinitive ending in –i corresponding to Italian –ere, as well as –ire.\(^{10}\)

(Bazzoni 1999: 57)

After proposing this two-class system, Bazzoni goes on to explain that there is another “category” of verbs, all of which are “irregular” and have an infinitive in –e. However, his motivation for leaving these verbs out of the two-class system is that they are irregular. After this explanation, he goes continues with a discussion of how there are really two types of first conjugation verbs (one of which involves the addition of a stem extension in certain paradigmatic slots, discussed below), as well as various other “irregularities” of the system. Thus, under the view proposed here, there are many more than two conjugational patterns of this variety of Italian (regardless of how many “classes” traditionally proposed).

Daco-Romanian, along with most other varieties of Balkan Romance, can be said to maintain the four-conjugational-class system of Latin. There are four different types of infinitives, each of which is expected from the four Latin starting points (as well the stress distinction in the 1/2PL of traditional 2\(^{nd}\)/3\(^{rd}\) conj. verbs discussed above). Most Balkan Romance grammars propose these four conjugational classes. One interesting exception particularly relevant here lies in the descriptions of the Daco-Romanian verb system given in Mallinson (1986), Juillard & Edwards (1971) among others. These

\(^{10}\)“Nel sassarese, a differenza dell’italiano, abbiamo soltanto due coniugazioni. Alla prima appartengono i verbi che all’infinito terminano in –a corrispondente all’italiano –are…Alla seconda coniugazione appartengono i verbi con l’infinito presente terminante in –i, corrispondente all’italiano –ere oppure –ire.”
sources propose five conjugational classes because there are actually five different infinitival endings in Romanian, the four described above, plus one in \(-i\), e.g., \(a \text{ hotări}^{i}\) ‘decide-INF’. This can be seen as controversial, as many consider this to be a subclass of the class in \(-i\) (e.g., Ruhlen 1974); some even propose that there are phonological rules that govern the distribution of \(i\) and \(i\) in these forms (Augerot 1974). However, these sources do not propose separate conjugational classes in other cases in which the system looks fractured to be discussed below. This, as far as I can tell, is because these causes of fragmentation do not affect the infinitive, while this specific case does. This just shows how pervasive the practice of classifying verbs by their infinitival vowel is in Romance (even more so than in Latin).

Note that those proposals that posit a two/three/four conjugational class system are not saying that not saying all verbs follow one of these two/three/four classes. Any argument that proposes such a small number of classes must rely on extensive subbranching and/or maintaining that irregular verbs are somehow outside of the system. However, there does not seem to be an advantage of having such a small number of classes if these classes are forced to have a complex internal structure in order to account for the different patterns observed.

3.2. Complicating Factors
The fact that there has been some leveling out of some aspects of Latin verbal morphology into Romance does not mean that the system is necessarily moving towards a “more simplified” system. Along with analogical change leading to what could be argued as a greater level of regularity, analogical change has also had the opposite effect – the fragmenting of the Romance conjugational system. However, this does not mean that linguists normally consider the number of Romance conjugational classes to be any higher than the traditionally defined four for Latin. In cases where something
has emerged to cause the fragmentation one of the traditional conjugational classes inherited from Latin, it is usually considered to be a "subclass" of the conjugation with which it shares a theme vowel.

3.2.1. Phonological complication

In some cases, phonology has complicated the picture. A case of this is discussed in Chapter 1 regarding verbs with alternating stems in Spanish. Recall that some verbs that have root vowels e and o show diphthongization in certain forms, while other verbs with the same root vowels do not. Since it is not synchronically predictable which verbs show stem diphthongization and which stems do not, it could be said that the Spanish verbs can be divided into two groups depending on if they show stem alternations or not (or, combining this factor with the traditional manner of classification, the verbs can be divided into even more groups). The fact is that this complication is caused (diachronically) by a series of regular sound changes\(^\text{11}\). Even something that is “regular” can have a widespread fragmenting effect. Similar phenomena are seen in other Romance languages, and specific cases will be described in detail in Sicilian and Aromanian in Chapter 6.

3.2.2. Morphological complication

There are several cases of complication in the evolution that can be more described as falling into the realm of morphology (or, in several cases “morpho-phonology”). There are several instances in which the adoption of additional material used in verbal stems has created fragmentation. For example, many Romance languages show a reflex of the

\(^{11}\) However, this really is not purely phonological, as there have been historical cases of verbs with a certain stem vowel not undergoing a certain (expected) evolution. Moreover, in certain dialects of Spanish, there has been shifting between diphthongizing and non-diphthongizing stems, e.g., in León, some speakers have *viendo* for Std. *vendo* 'sell-1SG' and *cerro* for Std. *cierro* 'open-1SG' (Zamora Vicente 1967). In Jalisco, some speakers have *ruempo* for Std. *rompo* 'break-1SG' and *quere* for Std. *quiere* 'want-3SG' (Cardenas 1967).
Latin affix [ESC] being used as a stem extension (see Chapter 4 for more on this). In most of the languages that show this is stem extension, it occurs solely in verbs of the language’s descendant of the traditionally described Latin 4th conjugation (Romance conjugations with theme vowel ĩ). In each of these languages, there are some verbs of the respective languages’ i-stem classes that do not show this. For example, the Italian verbs capire ‘understand’ and dormire ‘sleep’ are both considered i-stem verbs (of the Italian 3rd or 4th conjugation, depending if one follows a 3- or 4-conjugation system), even though capire shows the stem extension (e.g., 1SG capisco) while dormire does not (e.g., 1SG dormo). In most accounts of these languages, the i-stem conjugation is traditionally split into two subclasses, one of which is comprised of verbs showing the stem extension, and the other which does not (e.g., Badia i Margarit 1994, Pérez Saldanya, et al. 2004, etc. for Catalan, Atanasov 1990, Kovačec 1984, etc. for Balkan Romance). However, this is not always the case, as several accounts of Daco–Romanian, e.g., Lombard & Gâdei (1981), among others, consider these traditionally-defined “subclasses” to have the status of “conjugational class”.

The traditional Latin 1st conjugation (verbs with infinitives in –ARE) is normally considered to be the most “regular” of the traditional Latin conjugations. Most verbs of the Latin 1st conjugation, as shown in the previous chapter, have the present, perfect and participial stems formed essentially from the same base. This reputation for regularity and stability has been inherited into the Romance languages. However, despite this reputation, the Romance outcomes of the Latin 1st conjugation have been subject to fragmentation. When looking at the standard, “core” Romance languages, this fragmentation is not particularly clear; it appears to a greater extent in the peripheral, less-studied members of the family.

---

12 Recall that there are several exceptions, particularly those verbs having a perfect in –ūī (rather than the more common –Avī)
For example, the Latin affix [ESC] has developed into a stem extension in some verbs of the Romansch 1st conjugation, e.g., *cuntinuar* ‘continue–INF’, 1SG *cuntinesch*). Another striking example is found in the development of another Latin affix (again, to be discussed in detail in Ch. 4) into a stem extension in some 1st conjugation verbs in most varieties of Balkan Romance, e.g., DR *lucrez* ‘work–1SG’ showing the affix vs. cânt ‘sing–1SG’ not showing it, as well as in some varieties of Italian, e.g., Corsican (Dalbera–Stefanaggi 1997), and Sassarese, e.g., *barièggiu* ‘tolerate–1SG’, cf. *barià* ‘tolerate–INF’, (Bazzoni 1999), among other varieties.

Another complicating factor that has emerged in the development of several of the Romance languages is the presence of an unetymological velar consonant in several paradigmatic slots, most typically the 1SG present indicative and throughout the present subjunctive paradigm, e.g., Sp. *tengo* ‘have–1SG.PRES.IND’, *tenga* ‘have–1SG.PRES.SUBJ’, cf. *tener* ‘have–INF’). While in Spanish and Italian the group of verbs showing this unetymological consonant is limited in number and mainly confined to the verbs with infinitives in –*er(e)* (essentially, the outcome of the Latin 2nd/3rd conjugations), in Catalan it has been extended to many more verbs without regard for traditional conjugational class boundaries. Since its presence in the lexicon is not necessarily predictable, these verbs are usually considered to be “irregular” in most accounts of Spanish and Italian. In Catalan, and especially non-standard dialects thereof, the population of verbs showing this unetymological velar consonant is so high that classifying all of these verbs as “irregular” seems no longer a viable option. This phenomenon, with emphasis on its distribution in non-standard dialects of Catalan, will be discussed in much further detail in Chapter 5.
In the Alguerese Catalan\textsuperscript{13}, the 1\textsuperscript{st} conjugation has been fragmented into three groups (Lloret 2009). One follows the general Latin pattern of conjugation, one has the stem extension descended from Lat. –\textsc{idi}–, and one shows the velar augment.

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|}
\hline
 & INF & 1SG.PRES.IND. \\
\hline
w/o stem extension & cantar & canto & ‘sing’ \\
with velar augment & canviar & canvic & ‘change’ \\
with reflex of –\textsc{idi}– & odiar & odietj & ‘hate’ \\
\hline
\end{tabular}
\caption{Verbs in –\textsc{a}, Alguerese Catalan (Lloret 1999)}
\end{table}

The development of the group with a stem extension in –\textsc{idi}– in this Catalan dialect is clearly due to language contact, as it occurs in the neighboring Sassarese dialects of Italian, while not occurring in any other dialect of Catalan. Regardless of its source, it now forms part of the verbal system of this dialect. Speakers do not care where this came from, but rather, only the fact that exists in these verbs.

3.3. \textbf{How many conjugational class does Romance really have?}

As discussed earlier, most descriptions of Romance conjugational classes propose somewhere between two and five conjugational classes. This number is generally completely based on how many different theme vowels infinitives in that language have. As these authoritative grammars do not ignore the verbs that diverge from any of these patterns, this requires subbranching of each of the conjugational classes, as well as often singling certain verbs out as “irregular”. For example, Wheeler et al. (1999) describe the Catalan 2\textsuperscript{nd} conjugation (verbs with infinitives in –\textsc{er}, –\textsc{re} and a few with no theme vowel) as having several subclasses, and nearly all descriptions of Romance languages showing the stem extensions from Latin [ESC] posit a split of the language's

\textsuperscript{13} Alguerese Catalan is spoken in the city of Alghero (Cat. L’Alguer) in northwestern Sardinia. The presence of this dialect is the vestige of Aragonese domination and colonization of much of the insular Mediterranean.
respective –i stem class into two subclasses based on the presence/absence of the stem extension.

As discussed in Chapter 1, Aguirre & Dressler (2006) give an in-depth analysis of the Spanish conjugational system based on the principles laid out in Dressler (2003). The system is first broken down into two macroclasses, which depend on the thematic vowel the verb shows (a for one microclass vs. e/i for the other). However, as Spanish verbs are obviously conjugated following more than two patterns, these macroclasses must be broken down into a series of subclasses. Their first macroclass consists of four microclasses as well as four verbs that follow independent patterns. The other macroclass is split into two "classes" (essentially a level of categorization between "macroclass" and "microclass"), an e-class and an i-class. The e-class has twelve microclasses and ten verbs that follow independent patterns and the i-class has eleven microclasses and eight verbs that follow an independent pattern. Adding these numbers up gives twenty-seven microclasses and an additional twenty-two verbs that follow patterns not followed by any other verb. Adding these gives a total of forty-nine patterns (a number that is still quite a bit less than that of Latin, once all the details are taken into account) to which we can add Spanish defective verbs, which will still increase this number.

However, an important question is raised (in the title of his article) by Sánchez Miret (2000) "¿A quién le importa cuántas conjugaciones hay en Español?". He argues that it does not really matter how many conjugational classes a language has, and instead of trying to come up with some concrete number, linguists should be focusing on answering questions about how speakers classify these verbs into conjugations.

The data show quite a bit of fragmentation (in Spanish, as well as in other languages), which I believe shows that speakers may have some core, fairly widespread
generalizations, but in the periphery, the generalizations are quite local and the verbs show quite a bit of lexically-specific properties as to the patterns in which they inflect.

3.4. **Summary and Conclusions**

In some ways, it can be said that the Romance conjugational system(s) are simplified when compared to that of Latin. However, this is balanced out with innovations in Romance as a whole, or in particular Romance languages, that have led to increased fragmentation in Romance when compared to Latin. Chapter 2 laid the groundwork for why this traditional system falls short for Latin, and this chapter shows while it is true that there has been some simplification in the evolution of Romance, the system, essentially still that used for Latin, falls just as short when describing Romance. Chapters 4 and 5 discuss two of the “complicating” factors briefly introduced in this chapter in detail.
CHAPTER 4:
ON THE VERBAL EXTENSIONS [ESC] and [EDZ] IN BALKAN ROMANCE

4.0. INTRODUCTION

This chapter discusses certain verbal conjugational patterns involving stem extensions in the Balkan Romance languages. While these patterns are typically considered to be subclasses of the traditional conjugational classes posited for Romance, I argue that this is not necessarily the case. After giving a brief history of these forms, traditional views of the Daco-Romanian conjugational class system and general issues regarding the inventories of these so-called “sub-classes” are discussed. This is followed by an examination of cases of historical, geographical and contemporary variation using a variety of data sources (e.g., old Daco-Romanian texts, dialect atlases, the internet, etc.). While Daco-Romanian is the main focus of this chapter, the situation in the other Balkan Romance languages is given some attention in the concluding sections.

While most traditional grammars of the Balkan Romance languages consider these patterns to be subclasses, I argue that there is no legitimate reason to do so. Instead, I argue that they are independent classes that are organized on the basis of multiple features. Allowing for a recognition of multiple features (rather than one overriding, primary feature, as in traditional views of Romance conjugational classes) being involved in verb categorization allows us to maintain the reference to the shared features that led these traditional subclasses to being traditionally described as such, as well as allowing to show other connections that help account for cases of variation and change.
4.1. PRELIMINARIES

4.1.1. The Balkan Romance Languages

Balkan Romance forms part of the Eastern group\(^1\) of the Romance language family, along with Dalmatian\(^2\) and Central/Southern varieties of Italian. The Balkan Romance group is composed of four separate languages, traditionally divided into two groups based on the location in which they are spoken with respect to the Danube River.

<table>
<thead>
<tr>
<th></th>
<th>Number of Speakers</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Danubian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daco–Romanian</td>
<td>~25 million</td>
<td>Romania/Moldova(^3)</td>
</tr>
<tr>
<td>Istro–Romanian</td>
<td>~1000 speakers</td>
<td>Istrian Peninsula</td>
</tr>
<tr>
<td>South Danubian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aromanian</td>
<td>~300,000 speakers</td>
<td>Greece/Albania/Serbia/Bulgaria/Macedonia/Romania</td>
</tr>
<tr>
<td>Megleno–Romanian</td>
<td>~10,000 speakers</td>
<td>Greece/Macedonia</td>
</tr>
</tbody>
</table>

Table 4.1: The Balkan Romance Languages

The Balkan Romance group's isolation from "Latinity" in general (e.g., geographical isolation from the other surviving descendants of Latin, cultural differences, etc.) and centuries of language contact and bilingualism as part of the Balkan Sprachbund, have given the Balkan Romance languages a distinct character when compared to the other members of the Romance family. After the split of their common ancestor (usually called "Proto-Balkan Romance" or "Common Romanian"), the different contact situation

---

\(^1\) This grouping is traditionally defined by a few isoglosses, specifically the loss of Latin final -s and the lack of voicing of intervocalic Latin stops, but it is not at all clear nor uncontroversial; see Tagliavini (1969), von Wartburg (1971) Malkiel (1978, 1991) for discussion on this issue.

\(^2\) A now–extinct Romance language (or languages) that was spoken along the Dalmatian Coast (Adriatic Coast of modern Croatia)

\(^3\) There has been some controversy as to the identity of the primary language spoken in Moldova. Prior to Moldova's independence from the Soviet Union in 1991, the official language was "Moldovan" and was written in the Cyrillic alphabet. The 1994 Moldovan constitution declares "Moldovan" written in the Latin alphabet to be the official language of the state: "Limba de stat a Republicii Moldova este limba moldovenească funcționînd pe baza grafei latine." The unrecognized breakaway republic of Transnistria considers "Moldovan" (written in the Cyrillic alphabet) to be co–official with Russian and Ukrainian. The controversy is political, and linguistically, Moldovan is unquestionably Daco–Romanian (Dyer 1996, 1999).
that in which each was involved has given each Balkan Romance language a distinct character as well.

Although the other Balkan Romance languages deserve attention, this chapter primarily focuses on standard Daco–Romanian (commonly also referred to as “Romanian”), the Balkan Romance language with by far the largest number of speakers and on which by far the greatest amount of research has been conducted. As much attention is given to non-standard Daco–Romanian varieties as possible, given the available resources. The last few sections of this chapter briefly discuss similar situations in the other Balkan Romance languages. See Chapter 6 for a much more detailed treatment, though not completely focused on this phenomenon, of a specific Aromanian dialect.

4.1.2. Stem extensions [ESC] and [EDZ]⁴ – From Latin to (Balkan) Romance

As briefly introduced in the previous chapter, certain Romance languages have inherited the Latin affixes –ISC–/–ESC– and –IDI–. However, in most of these languages, they have lost the original function and are used as stem extensions in certain paradigm slots in certain inflectional classes.

4.1.2.1. [ESC]

Latin had a suffix –ISC–/–ESC– with two (somewhat similar) major functions. When it was attached to a verbal form, the suffix had an inceptive or inchoative function:

\[ \text{VERB} + \text{[ESC]} \rightarrow \text{”to start to VERB” or ”to begin VERBing”} \]

\[ \text{e.g., Lat. DORMIŌ ‘sleep–1SG’ + [ESC] } \rightarrow \text{ DORMIISCŌ ‘fall asleep–1SG’} \]

⁴ I will call these affixes [ESC] and [EDZ] throughout this dissertation, regardless of the language discussed. They do have different outcomes in different Romance languages, but I use this terminology for convenience.
When attached to noun or adjective roots, the affix appears to have had a denominative/deadjectival function:

\[
\text{NOUN} + [\text{ESC}] \rightarrow \text{“become NOUN” or “do something prototypical with NOUN”}
\]

\[
\text{ADJ} + [\text{ESC}] \rightarrow \text{“make ADJ” or “turn ADJ”}
\]

e.g., Lat. PULCHER ‘beautiful’ + [ESC] \(\rightarrow\) PULCHRESČō ‘make beautiful–1SG’

A limited number of verbs was carried from Latin into Romance with the affix already attached\(^5\). In this small number of cases, there is no paradigmatic variation; it appears in nearly\(^6\) all forms, regardless of person, number, tense, mood, etc. These lexical items are consistently found across Romance with the affix attached, e.g., Lat. COGNOSCō ‘know–1SG’ gives Ptg. conheço, Sp. conozco, Fr. je connais, It. conosco, DR cunośc, etc.

Besides these few cases, the different Romance languages vary in their treatment of this affix, both semantically and morphologically. Lausberg (1962) explains that the Romance family can be divided into two groups depending on their treatment of this Latin affix. In Sardinian, Spanish and Portuguese, [ESC] maintains the function/semantics that it had in Latin, e.g., Sp. florecer ‘to flower’, cf. flor ‘flower (noun)’; Ptg. anoitecer ‘to become dark’, cf. noite ‘night’). While in Latin, the suffix was nearly almost restricted to the forms within the present system, the suffix has been extended throughout out the paradigm in these Romance languages (in a similar development to that seen in the Latin verb POSČō ‘ask’, discussed in Chapter 2).

In the rest of Romance\(^7\) (Catalan, Gallo–Romance, Rhaeto–Romance, Italo–Romance, Balkan Romance), while the inchoative/inceptive meaning of the affix has

---

\(^5\) This is similar to how a few verbs were carried into Latin from PIE with this affix already attached; see Chapter 2.

\(^6\) There are some exceptions, e.g., the Italian perfect of this verb, that are not discussed here.

\(^7\) Bartoli (1898) claims that the Ragusan variety of Dalmatian (spoken around Dubrovnik, extinct since about 1500) joins this group, while the Vegliot dialect does not. Due to the very limited nature of available Dalmatian data, these claims must be taken with some caution.
been lost, a reflex of it has developed into a verbal stem extension in certain verbal forms. It is found most notably, but not exclusively, in some verbs of the languages’ respective i-stem classes (the descendant of the traditional Latin 4th conjugation, e.g., Fr./Cat. verbs in –ir, lt. verbs in –ire, DR verbs in –i, etc.). Examples given below:

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>1SG.PRES.IND.ACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat.</td>
<td>llegir</td>
<td>llegeixo</td>
</tr>
<tr>
<td>Fr.</td>
<td>finir</td>
<td>je finis</td>
</tr>
<tr>
<td>Rms.</td>
<td>referir</td>
<td>referesch</td>
</tr>
<tr>
<td>Frl.</td>
<td>partì</td>
<td>partis</td>
</tr>
<tr>
<td>lt.</td>
<td>capire</td>
<td>capisco</td>
</tr>
<tr>
<td>DR</td>
<td>a vorbi</td>
<td>vorbesch</td>
</tr>
</tbody>
</table>

Table 4.2: [ESC] as stem extensions in i-stems throughout Romance

While the specific verbs that show the stem extension vary across the languages in question, the paradigmatic distribution shows some consistency. In Italian, Catalan, Rhaeto–Romance, and throughout Balkan Romance, the stem extension appears in the singular and 3PL forms in both the present indicative and subjunctive paradigms. In Standard French, as well as in some Italian and Friulian dialects, Rohlfs 1966, Frau 1984, the verbal extension has been extended to a further extent, for example, French 1/2PL present indicative forms (e.g., vous finissez ‘finish–2PL.PRES’), as well as the imperfect (e.g., je finissais ‘finish–1SG.IMP’). While it is confined to the i-stem class in most of these languages, this is not the case in all of Romance, as it appears frequently in both the a-stems and i-stems of some Romansh varieties of Rhaeto–Romance, e.g., Surselvan a-stem gratular ‘greet–INF’, gratuléeschel ‘greet–1SG’ (Lausberg 1962).

In respect to the lexical items associated with the affix across the languages, there is a substantial amount of variation. The presence of the stem extension in a verb in one language does not necessarily mean that its cognates in the other languages will

---

8 There are differences in vowel quality, as some languages have a outcome descended from *isc (e.g., Italian, French) while others from *esc (e.g., Catalan, Balkan Romance)
show it as well. For example, from Lat. SERVĪO 'serve–1SG', Std. French has 1SG je sers (rather than je servis) and Std. Italian has servo (rather than servisco), both cases without the extension. However, Std. Catalan has 1SG serveixo, and Megleno–Romanian\(^9\) has 1SG sirbéś (Capidan 1935), both showing the stem extension.

One preliminary consistency seen is that in the reflexes of some ‘core’ verbs of the Latin 4\(^{th}\) conjugation, the stem extension is not present in any of the languages in question. For example, reflexes of Lat. DORMĪŌ 'sleep–1SG' include Cat/It. dormo, Fr. je dors, Rms./DR dorm, all without the stem extension. However, when looking outside of the standard varieties of the “major” Romance languages, this generalization does not hold. For example, some varieties of Catalan show reflexes of Lat. DORMĪŌ with the stem extension, e.g., [dormisko], [durmiski] (Alcover & Moll 1929, Griera 1966, Viaplana, et al. 2007) as well as other verbs normally not conjugated with the affix, e.g., Lat. *moriō \(\rightarrow\) Std. Cat. moro, Aranese [muriski] ‘die–1SG’ (Renat i Ferris 1933, Winklemann 1990).

In most traditional grammars of the languages that show a reflex of Latin [ESC] as a stem extension in some verbs in their respective i-stem conjugation, this conjugation is split into two subclasses, one showing the stem extension, and the other not showing it. Independent conjugational class status is rarely given to these traditional “subclasses”.

4.1.2.2. [EDZ]

The Latin suffix –IDI–, much like the affix discussed above, had a denominative function. Latin –IDI– gives Balkan Romance –edz which is maintained in varieties dialects of Aromanian, and becomes –ez in most dialects of Daco–Romanian and Megleno–Romanian (Caragiu–Marioteanu 1975). This has the same paradigmatic distribution in

\(^9\) Megleno–Romanian sirbéś shows the expected evolution from Latin SERVĪO, as well as a reflex of the [ESC] affix. Note that the Daco–Romanian verb a servi (1SG servesc) is definitely a later borrowing (from French or Italian). If this were a native Daco–Romanian word, something like a sârbi would be expected, due to the Common Romanian \(v \rightarrow b\) and \(e \rightarrow ā\) changes.
Balkan Romance as the [ESC] suffix (i.e. in the present indicative/subjunctive singular and 3PL) but it is restricted to a-stem verbs, some of which show it (e.g., DR *lucrez* 'work–1SG' and some of which do not (e.g., DR *cânt* 'sing–1SG').

While not as widespread in Romance as the [ESC] infix, this affix (with this distribution/function) is not limited to Balkan–Romance,\(^{10}\) as it is also present in some Northern Italian dialects, e.g., Corsican and Northern Sardinian, and some Southern Italian dialects, e.g., Puglia and Lucania (Lausberg 1962), and the Catalan dialect of Alghero briefly discussed at the end of Chapter 3. This also had an apparently larger distribution in the now–extinct Vegliot variety of Dalmatian\(^{11}\) (Ive 1886, Bartoli 1906).

Again, as is the case with the affix [ESC] discussed above, most traditional grammars of the Balkan Romance languages (as well as the few other dialects that show it) consider the 1\(^{st}\) (or a-stem) conjugation to be broken down into two subclasses; one made up of the verbs that show the stem extension [EDZ] and the other that is made up of the verbs that do not. Again, these individual “subclasses” are rarely given status as full conjugational classes.

### 4.2. DACO–ROMANIAN CONJUGATIONAL CLASSES

#### 4.2.1. Traditional views

“Traditional” views of the Daco–Romanian (as well as those of the other Balkan Romance languages) system typically follow the Latin system when describing conjugational classes, thus, as in the traditional accounts of Latin, a difference in theme vowels, particularly in the infinitival forms, equals a difference in conjugational class affiliation.

Some descriptions (e.g., Tiktin 1905, Rauta 1947) list four conjugations, one

---

\(^{10}\) It is not necessarily even found across all of Balkan Romance, as it appears to be absent from Istro–Romanian (as far as I can tell from the available sources). See section 4.5.3.

\(^{11}\) Interestingly enough, while this started in a-stem verbs in the languages discussed above (e.g., Veg. *kantaja* 'sing–1SG' < Lat. *CANTO*, where –aj is the reflex of Lat. –IDI) it spread to other stem types in Vegliot, e.g., Veg. *potaja* 'be able–1SG' < Lat. *poteó* (e-stem); Veg. *venaja* 'come–1SG' < Lat. *VENIÔ* (i-stem) (Bartoli 2000).
corresponding to each of the Latin conjugations. This is not to say that the inventories
of each has been maintained, as there has been the loss of Latin verbs, the addition of
non-Latin verbs, and shifts in traditional class membership.

<table>
<thead>
<tr>
<th>Lat.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-ĀRE</td>
<td>-ĒRE</td>
<td>-ĒRE</td>
<td>-ĪRE</td>
</tr>
<tr>
<td>DR</td>
<td>-a</td>
<td>-ea</td>
<td>-e</td>
<td>-i</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.3: Evolution of DR infinitives (take 1)

There is comparatively small group of verbs with infinitives (and a distinct conjugational
pattern) in -ī, which some sources (e.g. those above) consider to be part of the 4th DR
conjugation, mainly due to their small number and being arguably phonologically
predictable (see Augerot 1974). Others (Mallinson 1986, Julliand & Edwards 1971)
consider there to be five DR conjugations, the four above, plus one made up of verbs
with an infinitive in -ī.

<table>
<thead>
<tr>
<th>Lat.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-ĀRE</td>
<td>-ĒRE</td>
<td>-ĒRE</td>
<td>-ĪRE</td>
</tr>
<tr>
<td>DR</td>
<td>-a</td>
<td>-ea</td>
<td>-e</td>
<td>-i</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.4: Evolution of DR infinitives (take 2)

According to the principles established in Chapter 1, the split of -ī and -ī verbs into
separate conjugations is definitely justified12. The question that must be raised, though,
is why is this "split" justified as seen in many descriptions, but other "splits" are rarely
seen as justified by traditional sources (though there are exceptions, as discussed

12 Infinitives in -ī are always preceded by r, e.g., coborī 'lower', urī 'abhor'. However, there are also
infinitives in -i preceded by r, e.g., muri 'die' urmāri 'follow'. It has been proposed that there is
an underlying geminated r in the verbs that have the ending -ī, and potential evidence is seen in
Latin verbs with a geminated r that develop into DR verbs in -ī, e.g., Lat. HORRIRE \(\rightarrow\) DR urī
(Augerot 1974). However, this analysis is too dependent on Latin and abstract phonology, so it is
best to consider them different patterns.
below). The answer probably lies in the fact that the verbs in \(-i\) and verbs in \(-i\) have different infinitives, while, say, the subgroups of verbs in \(-a\) that show or do not show a reflex of the stem extension [EDZ] both still have an infinitive in \(-a\). I argue, on the other hand, that the infinitive does not really matter; while there are arguments for classifying them together, there are also convincing arguments for having them form independent conjugational patterns.

As is the case in the other Romance languages discussed in the previous section, the common view is to divide the class of Balkan Romance \(i\)-stem verbs (as well as the class of \(i\)-stem verbs) into two sub-classes: (a) verbs that have a stem extension [ESC] in various slots in the paradigm\(^{14}\) (the [+ESC] subclass) and (b) those that do not (the [-ESC] subclass); see examples of each of the subclasses below:

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRES.IND.</td>
<td>doresc</td>
<td>dorești(^{15})</td>
<td>dorește</td>
<td>dorim</td>
<td>doriți</td>
<td>doresc</td>
</tr>
<tr>
<td>PRES.SUBJ.</td>
<td>doresc</td>
<td>dorești</td>
<td>dorească</td>
<td>dorim</td>
<td>doriți</td>
<td>dorească</td>
</tr>
</tbody>
</table>

Table 4.5: PRES.IND./SUBJ. paradigms of DR a dori 'want' [+ESC]

\(^{13}\) In addition, the traditional 2\(^{nd}\) (verbs in \(-ea\)) and 3\(^{rd}\) (verbs in \(-e\)) are conjugated in the present tense nearly identically, as only the 1PL/2PL show a difference (the stress distinction discussed above). On the other hand, the traditional subclasses of the 1\(^{st}\) (or 4\(^{th}\) conjugation) show more differences in the present system. However, the vast majority of DR sources do not propose that the 2\(^{nd}\) and the 3\(^{rd}\) are 'subclasses' of a single conjugation. One argument for this point of view can be seen in Ruhlen (1974) and evaluated in a response by Juillard (1978). In addition, I have seen sources that claim this for Aromanian (though in these dialects, there may have been an actual collapse of the traditional 2\(^{nd}\) and 3\(^{rd}\) conjugations, at least in the present tense). I definitely agree that they should be separated because there are differences, though I do not think traditional Romanian grammarians had the same motivation (otherwise the traditional subclasses of the 1\(^{st}\) conjugation would be considered independent classes too!). Though I completely disagree with Ruhlen’s point of view, he is at least consistent in his definition of what he believes to be a class.

\(^{14}\) Again, as discussed earlier, as in the other standard Romance languages that show this stem extension (except for French), it appears in the singular + 3PL of the present indicative. It also appears in the same slots in the present subjunctive, along with the 2SG imperative (which is the same forms as the 3SG.PRES.IND). Outside of the standard language, it can be found elsewhere as well, e.g. infinitives, 1/2PL.PRES forms, etc.

\(^{15}\) DR <esc> [esk] is palatalized to <eșt> [eʃt] before front vowels (2SG/3SG indicative, 2SG subjunctive forms).
As for the pseudo–conjugation in –i, some members show a reflex of Latin [ESC] with the same distribution, but it shows up as –așc, e.g., a hotărî `decide-INF', 1SG hotărășc. Thus, just like the verbs in –i, this class must also be broken down into two subclasses if we are going to take it as legitimate.

Similar to how the –i–stem verbs (as well as the –i–stem verbs) are traditionally divided into two groups depending on the presence or absence of the extension [ESC] (or [ĂSC]), the a–stem verbs are also divided traditionally into two groups depending on if they show a reflex of Latin –iDI– as a stem extension ([+EDZ] subclass), or if they do not ([–EDZ]) subclass. Examples below:

Table 4.7: PRES.IND./SUBJ. paradigms of DR a lucra 'work' [+EDZ]

Table 4.8: PRES.IND./SUBJ. paradigms of DR a cânta 'sing' [–EDZ]

Given that these are normally considered subclasses, the following system (ignoring all other oddities in Romanian verbal morphology) is typical of descriptions of the DR conjugational system.
4.2.2. Less traditional views

While descriptions such as those above are typical, it is essential to note that there are alternative ways of looking at the Daco-Romanian (or Balkan Romance as a whole) verb classes that are (somewhat) more in line with the views suggested by this dissertation. Lombard (1974) and Lombard & Gâdei (1981), propose six conjugations, where the two traditional subclasses of the traditional 1st and 4th conjugations are considered to be independent conjugations. Also notable is the fact that Lombard has the i-stems and î-stems collapsed, which as discussed above, is not completely predictable distinction.

Verbal inflection includes six conjugations, that is to say, six systems of endings:

I: Infinitive in –á, and the 1SG.PRES.IND. with no ending, e.g., jurá ‘swear-INF’, jur ‘swear-1SG’

II: Infinitive in –á and the 1SG.PRES.IND. in –éz, e.g., formá ‘form-INF’, forméz ‘form-1SG’

III: Infinitive in –í and the 1SG.PRES.IND. with no ending, e.g., fugí ‘run-INF’, fug ‘run-1SG’

IV: Infinitive in –í and the 1SG.PRES.IND. in –ésc, e.g., iubi ‘love-INF’, iubesc ‘love-1SG’

V: Infinitive in –eá, e.g., tăcea ‘silence’
VI: Infinitive in \(-e\) (unstressed), e.g., \(tême\) 'in the reflexive verb \([a]\ se teme\ 'fear'

(Lombard 1974:28–29)\(^{16}\)

Other sources are even more explicit in defining verb classes based on how they actually conjugate, rather than following the tradition inherited from Latin. Ciompec, et al. (1985) propose ten Romanian conjugations, while Felix (1964) proposes twelve. Unlike the other attempts described above, their classification is not exclusively based on the present tense. That is, the perfect tense, the past participles, the gerund, etc., which show some differences within the traditional 4/5/6-conjugational classes, are taken into account.

Most significant to the issues brought up by this dissertation is Juilland & Edwards (1971), which, while proposing only five conjugations, has the goal of identifying features that the traditional conjugations have in common with other conjugation. This proposal, as it is definitely relevant to the discussion here, is evaluated in further detail at the end of this chapter.

The remaining portion of this chapter discusses these traditionally-defined 'subclasses' exclusively. There are other oddities of the Balkan Romance verb (e.g. issues in the perfect system, past participles formation, etc.) but these are left to others to discuss (see Chapter 6 for a discussion of Aromanian conjugation). In what follows, the inventories and historical development of each of these classes are be discussed with specific regard to Daco–Romanian (though attention is given to the rest of Balkan Romance at the end).

\(^{16}\) La flexion du verbe comprend six conjugaisons, c'est-à-dire six systèmes de terminaisons (de désinences): I: l'inf. a \(-á\), et le prés. ind. 1 la term. zéro, ex. jurá 'jurer', jur 'je jure'; II: l'inf. a \(-á\) et le prés ind. 1 \(-éz\), ex. formá 'former', forméz 'je forme'; III: l'inf a \(-í\) et le prés. ind. 1 la term. zéro, ex. fugí 'courir' fug 'je cours'; IV: l'inf. a \(-í\) et le prés ind. 1 \(-ésc\), ex. iubi 'aimer', iubesc 'je aime'; V: l'inf. a \(-eá\), ex. tâcea 'se taire'; VI: l'inf. a \(-e\), ex \(tême\), dans le verbe réfléchi \([a]\ se teme\ 'craindre'\)
One of the tenets of the Gross Detail Method described in Chapter 1 is that the history of the development of a language does not matter to the speaker of that language. This is true, though history does matter to the linguist. Here I make no claims that much of the following information really “matters” to a native speaker of Daco–Romanian. However, the way in which these verb classes developed over time show us the generalizations of speakers in action. Following the discussion of the formation of these classes, a more in-depth discussion of the synchronic status of these classes is presented, which shows that essentially the same actions that, over time, led to the formation of these patterns, are occurring synchronically as well.

4.3. INVENTORIES AND DEVELOPMENT OF PATTERNS IN DACO–ROMANIAN

4.3.1. Development of the $i$–stem class(es), preliminary generalizations

The group of verbs following the $i$–stem/$[-\text{ESC}]$ pattern is primarily composed of core vocabulary inherited from Latin (similar to the Pan–Romance 'trend' discussed in 4.2., above).

<table>
<thead>
<tr>
<th>Lat.$-1\text{SG}$</th>
<th>DR.$-1\text{SG}$</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audiō</td>
<td>aud</td>
<td>'hear'</td>
</tr>
<tr>
<td>Dormiō</td>
<td>dorm</td>
<td>‘sleep’</td>
</tr>
<tr>
<td>Fugiō</td>
<td>fug</td>
<td>‘flee’</td>
</tr>
<tr>
<td>Sciō</td>
<td>ştiu</td>
<td>‘know’</td>
</tr>
<tr>
<td>Veniō</td>
<td>vin</td>
<td>'come’</td>
</tr>
</tbody>
</table>

Table 4.9: Inherited Latin verbs , $-i/[-\text{ESC}]$

A preliminary view of the $[+\text{ESC}]$ group of the DR $i$–stems shows that it is heavily populated with verbs of foreign origin. This is particularly true of borrowings from other Balkan languages, most notably Slavic languages, including Old Church Slavonic, the liturgical language of much of the Balkans north of Greece (including Romania until the 19th century), and the dialects that were developing into Bulgarian and Serbian to the
south and west, respectively. There are also several Daco–Romanian verbs of Greek and Hungarian origin. The vast majority of borrowings from other Balkan languages entered DR following the *i*-stem/ [+ESC] pattern.

<table>
<thead>
<tr>
<th>source</th>
<th>1SG</th>
<th>DR–1SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCS¹⁷</td>
<td>čitajǫ ‘read’</td>
<td>citesc ‘read’</td>
</tr>
<tr>
<td></td>
<td>ljubljǫ ‘love’</td>
<td>iubesc ‘love’</td>
</tr>
<tr>
<td></td>
<td>pohošto ‘want’</td>
<td>poftesc ‘wish’</td>
</tr>
<tr>
<td></td>
<td>hrno ‘feed’</td>
<td>hrănesc ‘nourish’</td>
</tr>
<tr>
<td>Grk.</td>
<td>sóno¹⁸ ‘reach’</td>
<td>soneksi ‘arrive’</td>
</tr>
<tr>
<td>Hng.</td>
<td>lakik¹⁹ ‘live’</td>
<td>locuiesc ‘live’</td>
</tr>
</tbody>
</table>

Table 4.10: Balkan verbal borrowings into DR, –i/[+ESC]

The naïve generalization that one could make at this point is that the [+ESC] group is composed of verbs of foreign origin, while the [–ESC] group is composed of inherited verbs. In addition, there is another large group of DR verbs of a different origin that also follow the DR *i*-stem/[+ESC] pattern. These are denominative/deadjectival verbs, most of which are of Latin origin, though there are also verbs derived from nouns or adjectives from a variety of other (mainly Balkan) languages.

<table>
<thead>
<tr>
<th>source</th>
<th>DR 1SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lat.</td>
<td>înmultesc ‘multiply’</td>
</tr>
<tr>
<td></td>
<td>limpeșesc ‘clarify’</td>
</tr>
<tr>
<td></td>
<td>înlocuiesc ‘exchange’</td>
</tr>
<tr>
<td></td>
<td>numesc ‘name’</td>
</tr>
<tr>
<td>Grk.</td>
<td>folosesc ‘use’</td>
</tr>
<tr>
<td>Trk.</td>
<td>cântăresc ‘ponder’</td>
</tr>
<tr>
<td>OCS</td>
<td>zâmbesc ‘smile’</td>
</tr>
<tr>
<td>Bg.</td>
<td>îmbolnăvesc ‘make sick’</td>
</tr>
</tbody>
</table>

Table 4.11: Denominative/deadjectival verbs of Latin/Balkan origin, –i/[+ESC]

¹⁷ Standing in for the specific form of Slavic with which Proto–Romanians came into contact (which likely was rather similar to OCS).
¹⁸ The aorist stem is what is actually borrowed in Greek borrowings (e.g., *sos*– in this case). This –s became productive in assimilating other borrowings (most notably those from Turkish, but also those of other origins) into other Balkan languages (Sandfeld 1930).
¹⁹ The etymological dictionaries say this is from the Hng. verb *lakni* ‘live-INF’ (cf. *lakik* ‘live–1SG’). However, this could just be from the Hng. noun *lak*, meaning ‘dwelling’.
One case is particularly interesting, as two DR verbs are derived from the same root. The verb *dor ‘hurt–1SG’* is inherited from the Latin verb *DOLÉO*, while the verb *doresc ‘want–1SG’, is a denominative verb from the Latin noun *DOLOR ‘pain’*. The Latin verb maintains its conjugational affiliation in the transition from Latin to Daco–Romanian (2nd conjugation; Lat. *DOLĒRE*, DR *a durea*), while the noun undergoes a denominative process in which it becomes a verb following the *i*-stem, [+ESC] pattern.

Daco–Romanian also has corresponding nouns/adjectives in many of these cases (*sfredel ‘drill’, bolnav ‘sick’, etc.*), inviting the inference that the noun was borrowed and then a verb was formed from it via a derivational process already active in the language. Even in cases where there is not a corresponding noun in Standard Contemporary Daco–Romanian (e.g., there appears to be no modern DR noun *zâmb* from OCS *zöb*), this does not necessarily mean that one never existed\(^{20}\).

At this point, it appears that the *i*-stem/ [+ESC] class has two major components: (a) borrowed verbs and (b) derived verbs (regardless of the origin of the noun/adjective). We can combine these two together in a group called something like ‘adapted verbs’, as in both cases, some adaptation was necessary (either adaptation of a foreign verb into the DR verb system, or the adaptation of a noun/adjective into the DR verb system). Clearly, both cases are following a specific pattern that was already present in the language. When speakers are confronted with “making a verb”, they do so on the basis of their past experience in “making verbs”, whether from foreign material or from material already present in DR.

---

\(^{20}\) In fact, Daco–Romanian has lost quite a bit of its vocabulary of Slavic origin (at least in the Standard language) due to the language’s “relatinification” in the past few centuries (Close 1974).
4.3.2. Exceptions to these generalizations

While the vast majority of verbs of Balkan origin entered the \(i\)-stem/\(+\text{ESC}\) class, this is not the case of all Balkan borrowings into DR. There are several instances of these verbs entering the \(i\)-stem/\(-\text{ESC}\) class in Standard DR.

<table>
<thead>
<tr>
<th>source</th>
<th>DR-INF</th>
<th>DR-1SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hng.</td>
<td>a mântui</td>
<td>mântui</td>
</tr>
<tr>
<td></td>
<td>'save'</td>
<td></td>
</tr>
<tr>
<td>OCS</td>
<td>a trebui</td>
<td>trebuie</td>
</tr>
<tr>
<td></td>
<td>'need to-3SG'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a omorî</td>
<td>omor</td>
</tr>
<tr>
<td></td>
<td>'kill'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a slobozi</td>
<td>slobod</td>
</tr>
<tr>
<td></td>
<td>'free'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a stârui</td>
<td>stârui</td>
</tr>
<tr>
<td></td>
<td>'linger upon'</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.12: Borrowings, \(-i/-(-\text{ESC})\)

While these verbs do not show the stem extension in Standard DR, it is important to note that several of the above verbs do show variation across DR dialects as to which pattern, \(-\text{ESC}\) or \(+\text{ESC}\) they follow. Some of this variation is even listed in dictionaries, although additional details are lacking. For example, in the *Dictionarul explicativ al limbi romane*, non-standard variants involving these suffixes are often listed; e.g., the entry for *a mântui* includes a note that the 1SG.PRES. has a variant \(+\text{ESC}\) form, mântuiesc, cf. Std. mântui.

In addition to some verbs entering \(i\)-stem/\(+\text{ESC}\) class, there are several examples of Balkan borrowings that entered the \(a\)-stem/\(+\text{EDZ}\) class discussed above.

<table>
<thead>
<tr>
<th>OCS</th>
<th>DR-INF</th>
<th>DR-1SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>trupъ</td>
<td>a întrupă</td>
<td>întrupe</td>
</tr>
<tr>
<td>'body'</td>
<td>'make concrete'</td>
<td></td>
</tr>
<tr>
<td>kajati se(^{21})</td>
<td>a câinea</td>
<td>câinez</td>
</tr>
<tr>
<td>'repent-\text{INF}'</td>
<td>'complain'</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.13: Balkan borrowings: \(a\)-stem/\(+\text{EDZ}\)

\(^{21}\) The present root of this verb does not account for the presence of a \(-n-\) in the Romanian form. This is likely from a participial form, or from a noun (e.g. Modern DR has a noun câintă 'repentance' from the same root (*pocaanie* is attested in the *Codicele Voronețean*, cf. Bg. pokajanie 'repentance')
This pattern (a-stem/ [+EDZ]) is otherwise highly populated with words of Latin origin, a
great number of which were borrowed from French or Italian, e.g., DR a dansa 'dance–
INF', dansez 'dance–1SG' < Fr. danser 'dance–INF'). It appears that borrowed verbs (as
well as newly formed verbs from already present material) have mainly followed two
patterns throughout the history of Daco–Romanian. Earlier borrowings appear to have
mainly followed the i-stem/ [+ESC] pattern, while more recent borrowings appear to
have mainly followed the a-stem/ [+EDZ] pattern. However, there are possibly multiple
factors influencing the patterns these words followed upon entering the language,
several of which will be discussed in the following section.

4.3.3. Factors
Several factors appear to be influencing the pattern chosen by speakers of Daco–
Romanian when incorporating foreign lexical items (or creating verbs from already
present material) into their verbal system; four are discussed briefly below:

4.3.3.1. Time period in which the borrowing occurred
The most significant appears to be the time period in which the verb was formed. As
discussed above, early borrowings (from other Balkan languages) tend to enter the i–
stem/ [+ESC] class, while later borrowings (mainly from French) tend to enter the a–
stem/ [+EDZ] class. Drawing a line would be arbitrary and pointless, as there are both
rather old [+EDZ] verbs (e.g., lucrez, attested with the affix in the earliest Daco–
Romanian documents) and recent [+ESC] verbs, as it is still active for some speakers in
incorporating modern loanwords. This is not a situation in which the [+ESC] pattern was
“switched off” while the [+EDZ] pattern was “switched on”. This is more indicative of a
situation in which the productivity of one pattern has appeared to dim somewhat, while
the other pattern has gained productivity.
4.3.3.2. Phonology of the borrowed word

The phonological structure of the borrowed word also appears to be important, especially in the case of later borrowings. French verbs in –er (i.e., the French 1st conjugation) tend to enter the [+EDZ] group of the Romanian 1st conjugation (e.g., DR *dansez* 'dance–1SG' < Fr. *danser* ‘dance-INF’), while French verbs in –ir tend to follow the DR i-stem/[+ESC] pattern (e.g., DR *construiesc* 'build–1SG' < Fr. *construire* 'build-INF’). In this case, it appears that maintaining the French conjugational class may have been somewhat important incorporating these verbs into DR. This is not necessarily surprising given the prestigious status of French in Romania. The correspondences between French conjugational classes and DR conjugational classes are clear and obvious. However, since the systems are so different, there would be no way to maintain conjugational class of borrowings into Daco–Romanian from any of the (non–Romance) Balkan languages, as is discussed briefly below.

4.3.3.3. Origin of the borrowing

In theory, the origin of the borrowing should not matter whatsoever. The different patterns seen with Balkan borrowings on one hand (i.e., normally i-stem/[+ESC]), and French/Italian verbs, on the other (i.e., normally a-stem/[+EDZ]), have more to do with the time period in which DR speakers came into contact with populations speaking these languages.

However, the effect that French borrowings had on DR appears to have been significant. A possibility is that prior to the period of heavy French influence in DR, both

---

22 It is important to note that many assumed borrowings from French did enter the [-ESC] group of the 4th conjugation. For example, verbs such as *a conveni* (1SG *convin*), *a preveni* (1SG *previn*), etc. are assumed to be of French origin, but enter the [-ESC] class. These cases are attributable to analogy from DR *a veni* 'come-INF'. *vin* 'come–1SG' or represent DR calques on the French *preverb+venir* structure (Sanchez–Miret 2006)
the [+EDZ] class of the first conjugation and the [+ESC] class of the fourth conjugation were productive, but the [+ESC] class much more so. French was a language of culture in Romania and has been credited with affecting the language on many levels (e.g., French influence may have revived the use of infinitival constructions in Daco-Romanian, Close 1974, Joseph 1983). When French verbs entered Daco-Romanian, there may have been a conscious attempt to maintain their conjugational affiliation, thus meaning that many of these verbs entered the a-stem/[+EDZ] pattern, instead of the historically more productive i-stem/[+ESC] pattern.

South Slavic (as well as other Balkan) origins appear not to have had a notable influence on the class the borrowing entered. If we assume that it was a verb, rather than a noun, that was actually borrowed, then it is clear that there is not a direct correspondence between Romance conjugational classes and OCS (or Greek, Hungarian, Turkish, etc.) conjugational classes, so it seems like this was not a factor and most verbs entered a single class. The fact remains that the conjugational systems of Slavic and Romance are different, and maintaining conjugational affiliation between two Romance languages is much more straightforward than somehow maintaining it between a Slavic language and a Romance language (which could be a reason why the great majority of Slavic borrowings entered a single Romanian verb class).

Perhaps this influx of French words into Daco-Romanian could have in turn "promoted" the a-stem/[+EDZ] pattern to the status of the predominant productive class. It is the class into which most recent words are incorporated (though this is not always true; see section 4.4.3). On the other hand, French did not have the same prestigious status in Aromanian and Megleno-Romanian as it did in Daco-Romanian.

---

23 In modern South Slavic, on the other hand, there are verb classes that look somewhat similar to Romance classes with regard to their formation (they have a clear theme vowel, etc.) So, perhaps with more contemporary borrowings from Bulgarian, Serbian or Macedonian, this could be a factor. However, the issue is that words of these origins are not being frequently borrowed into Standard DR, the language to which we have the most access. Research on bilingual communities could show interesting trends.
So, if French words were not entering these languages, then there was nothing to spur on the increased productivity of the [+EDZ] class, which is a trend seen in MR and (though to a lesser extent) AR data. For example, some DR words of French origin in Romanian such as a dansa are not found in MR or AR (as far as I can tell). Another verb claimed to be of French (or Italian) origin in Daco–Romanian is a fuma ‘smoke’, 1SG fumez. AR has 1SG afum and MR has 1SG fum, both of which are claimed to be inherited from Latin FUMÆRE (Capidan 1935, Papahagi 1963).

4.4. VARIATION IN DACO–ROMANIAN

As discussed above, it appears as that there was a gradual transition in forming/adapting verbs from the i–stem/+ESC pattern to the a–stem/+EDZ pattern. If this is true, then we would expect variation. There is widespread variation when it comes to these verb–formation techniques. In addition to diachronic variation (that is, the difference regarding these patterns between DR of different eras), there is also geographic variation, as well as synchronic variation seen in the incorporation of recent loanwords into Daco–Romanian. This all shows us that what is going on today is the same thing that has been going on in the past; the building of the groups of verbs that follow these patterns occurred form–by–form.

4.4.1. Variation noted in grammars

Several grammars cite several cases of variation (sometimes prescriptively referred to as “errors”). For example, Niculescu (1981) notes that some Daco–Romanian speakers are leveling the distinction between the two subclasses of both the 1st and 4th conjugations in favor of the [-EDZ] and [-ESC] classes.

ARLR (1963) and Niculescu (1981) (among others) note that there are verbs that follow both of the i–stem models ([+ESC] and [-ESC]), as well as verbs that follow both
of the a-stem models ([+EDZ] and [-EDZ]). While these are often referred to as “the result of confusion” in grammars, all that has happened is that different speakers have incorporated these verbs into the system following different conjugational models. For example, the verb ‘prefer’ has a standard form a prefera as well as a “regional form” a preferi (1SG preferesc). Additionally, while the 1SG form of the standard is prefer (without a stem extension), some speakers have preferez (with the stem extension).

Interestingly, some of these cases of “double forms” have taken on different meanings. For example, there are two DR verbs derived from OCS ẓěbъ ‘tooth’: a zâmbi (i-stem/ [+ESC], 1SG zâmbesc) ‘smile’ and a zâmba (a-stem/ [+EDZ], 1SG zâmbez) ‘threaten, show teeth’. One could perhaps claim that a zâmbea is a later borrowing (perhaps from the Bulgarian verb zâbja se, which has a similar meaning.) However, this is not likely as Bulgarian (as well as Serbian and Ukranian) lost the earlier nasal vowels at some point in their development, and the nasal is preserved in the DR verb. The more likely is that there was a noun *zâmb present in Daco-Romanian (that has since been lost), that was verbalized twice, perhaps at different times. Another possible situation is that these could have been variants having roughly the same meaning, and then the present meanings were conventionalized.

Examples of variation are not limited to phenomena classified as “errors” in grammars. In the following sections, additional cases of variation within Daco-Romanian are discussed, including a brief glance at old Romanian texts, cases of geographical variation found in dialect atlases, and contemporary examples of variation found on the internet.

---

24 While lost in all of modern Slavic except for Polish and a few South Slavic dialects, the nasality of the Common Slavic vowels ę/ǫ is preserved in OCS loanwords into Daco–Romanian (realized as /ân/ or /un/ in the case of OCS ǫ) (Rosetti 1962).
4.4.2. Variation in Old Daco–Romanian texts

Compared to that of the rest of Romance, the first attestations of any Balkan Romance language come quite late. The oldest preserved document written in Romanian is the Scrisoarea lui Neacșu written in 1521 (compared to the 9th century for the earliest surviving French document, Tagliavini 1963), which unfortunately (for our purposes here) does not contain a single [+ESC] or [+EDZ] form. Codicele Voronețean, a text from the 16th century from Northern Romania does show these forms quite frequently. However, there are some differences in terms of their lexical distribution when compared to Modern Daco–Romanian. For example, there are a few examples in the Codicele Voronețean of verbs that are i-stem/ESC in Modern DR, but appear to be [−ESC] in the text. For example, the 3SG.PRES.SUBJ. form se rrapă is attested, from the verb a răpi ‘ravish’, whereas in Modern Standard DR, the form would be rapească (Costinescu 1981).

More numerous are attested forms of verbs in a that are +EDZ in Modern Standard DR but appear as [−EDZ] in Codicele Voronețean. For example, it shows 3SG.PRES.SUBJ. se veaghie where the corresponding form in Standard Modern DR would be veghează. In general, the [EDZ] suffix is comparatively rare in the Codice Voronețean. I have only come across four forms: two cases of the 3SG.PRES.IND.REF of a derepta, derepteadză; one case of the 1SG.PRES.IND. of a lucră, lucreadză; and one case of the 3PL.PRES.SUBJ. of a vâsla, se vânsleadză. Of course, we cannot be sure of the extent of the [EDZ] suffix in the language of the time (or location) of the Codicele Voronețean, as there could have been many other [+EDZ] verbs that are not attested in

25 The Voroneț monastery is in the north, close to the modern Ukrainian border
26 It appears to have still been /edz/ (rather than /ez/) at this time, or at least written with the Old Romanian Cyrillic letter с = [dz], rather than с = [z]
27 This really looks like a Modern DR 3SG. or 3PL.PRES.IND form of a lucră, lucrează, but Costinescu (1981) claims that it is a 1SG indicative form due to context.
28 Where this <n> came from in this form is unclear. This verb does not appear to be in Modern DR. There is a verb a vâsli in Modern DR; it is not clear if they are completely different verbs or if some sort of conjugational shift occurred.
the document, or even verbs that were commonly [+EDZ] could be attested, just not in forms that we would expect to see the [EDZ] affix (e.g. perfects, 1PL/2PL present forms, etc.).

There is no obvious variation within the Codicele Voronețean itself in regard to the presence/absence of these suffixes, which is not really surprising, as it was composed by a single author. However, looking across several texts could show that there was variation in this period (as there is today). In respect to the situation in early DR, Vasiliu, et al. (1986) claim that "regarding the inventory of verbs of the 1st and 4th conjugations which attach to the present the suffixes –ez, and respectively, –esc, have been showing differences between epochs as well as regions".

So, to really figure what the extent of these two affixes was, a more complete view of old Romanian texts is necessary. For example, there appear to be a few examples of variation in Codice Todorescu (Drăganu 1914), another Daco–Romanian text from the 16th century, though further investigation is warranted.

4.4.3. Regional variation

Much of the variation seen here is the result of contact with other languages, as borrowings are among the major sources of verbs that follow these patterns. This does not only exist between the Balkan Romance languages (as is discussed in 4.5., below), but also within Daco–Romanian as well. Daco–Romanian is spoken outside the modern borders of Romania, and even within Romania, speakers in certain areas come into contact with different languages than speakers living in other areas. For example, Daco–Romanian is also the principal language spoken in the modern Republic of Moldova. As Moldova was part of the Soviet Union between 1944 and 1991, Russian has had a considerable influence (not just limited to lexical borrowing) on the Daco–Romanian varieties spoken there (Dyer 1999). Dyer claims that Russian verbs entering
Moldovan varieties during the Soviet period entered the [+ESC] class of the traditional 4th conjugation (similar to earlier Slavic borrowings, but not more modern borrowings); e.g., Rus. *buksovati* ‘skid-INF’ gives Moldovan DR *bucsuiește* 'skid–3SG'.

In addition, there are varieties of DR spoken in the Serbian autonomous province of Vojvodina (discussed in Flora 1962 in the case of the variety of the municipality of Vršac), which likely have been influenced by Serbian. Also, within the borders of Romania, speakers come into contact with speakers of Ukrainian (in the northeast), Hungarian (in the northwest), Serbian (in the east) and Bulgarian (in the south). Plus, speakers at university likely have contact French (substantially in the past few centuries, but not as much recently) and English (ever increasingly).

The various Romanian dialect atlas projects undertaken throughout the 20th century frequently show interesting cases of variation regarding the presence/absence of the verbal extensions discussed here. *Atlasul lingvistic român* (1956) covers the entire country (along with one form each for the other Balkan Romance languages, which are unfortunately left blank in many of the maps). Along with adjectives and nouns, a few verbal forms were mapped, several of which show variation between the a-stem/[+EDZ] and i-stem/[+ESC] patterns. For example, in the case of *gunoiesc*29 ‘waste away, cover with manure–1SG’, while the majority of speakers have a 1SG form that is i-stem/[+ESC], two speakers (one from the center and one from the southeast) have a form following the a-stem/[+EDZ] pattern, *gunoiez*. A rather more dramatic example of variation is seen in the map for standard *rostează*30 ‘find, procure–3SG’. In this case, most northern speakers have the standard form, while there are several central speakers with a form with [ESC], *rostește*.

*Atlasul lingvistic român pe regiuni* (1960s–) is a more ambitious dialect atlas project that divides the country into seven regions roughly based on historical regional

---

29 Volume I, harta 13
30 Volume II, harta 482
boundaries. While the regional atlases vary in terms of size and of availability, interesting cases of variation are found throughout. First, there are cases of variation between $i$-stem [+ESC] and [-ESC] patterns. That is, while some speakers have categorized a verb as being [+ESC], others have categorized it as being [-ESC]. Examples of this are seen in the Transylvanian dialect atlas (see Costanzo 2008 for more details). While there are some reports that speakers are leveling the $i$-stem verbs in favor of the [-ESC] pattern (Niculescu 1981), a trend is seen in Northern Transylvania that shows that the opposite may be occurring. In three maps verbs that have standard DR [-ESC] verb forms, a quite large area of Northern Transylvania consistently shows [+ESC] forms. The areas in which the [+ESC] forms occur across the maps are not identical, but it appears that the extent is somewhat similar in the three cases. The fact that the areas are not identical demonstrates that this process is occurring form by form via local generalizations; it is more complete for some speakers, and less complete for others.

This set of atlases also shows interesting cases of $a$-stem [+EDZ]/[-EDZ] variation. For example, for lucrez\textsuperscript{31} 'work-1SG', several Oltenian speakers have lucru, without the affix. This is similar to the form found in Istro-Romanian (discussed in section 4.5.3.). This verb is one of the prototypical cases of a [+EDZ] verb; it is attested early (e.g., it was one of the few $a$-stem/[+EDZ] verbs attested in the Codice Voronețean), and appears in Aromanian and Megleno-Romanian as [+EDZ] (AR lúcredz, MR lucrédz). Interestingly, there are several other cases of [+EDZ] verbs in Standard DR that appear to be [-EDZ] in Northern and Northwest Oltenia. Additionally, there are a few Standard DR [+ESC] verbs that show the same trend in the same areas. For example, Std. DR cinez 'dine-1SG' and prânzesc 'have lunch-1SG' (among other examples) show forms without the affixes, cin and prînd, in similar geographic

\textsuperscript{31} \textit{Noul Atlas lingvistic român pe regiuni: Oltenia}, lucrez: harta 961, cinez: harta 278, prânzesc: harta 273
distribution (also similar, but larger, than the areas that show *lucru* for Std. *lucrez*). It would seem that there is an ongoing process effecting the loss of the suffixes in these parts of Oltenia (echoing the statement found in grammars that for some speakers, neutralization is occurring towards the [-EDZ]/[-ESC] classes).

The map covering verb *asurzesc*[^32] 'go deaf-1SG' in Muntenia și Dobrogea (most of modern southern Romania) shows several interesting cases of variation. While the vast majority of the forms found are *i*-stem/[+ESC], there are a few cases of a form with [EDZ] (*surdez, surzez*) as well of a few cases of forms without [ESC] or [EDZ], e.g. *asurd* (the absence of the suffix makes it impossible to tell if this is an *a*-stem or an *i*-stem).

The most dramatic example of variation I have come across in dialect atlases involves variation between the *i*-stem/[+ESC] and *a*-stem/[+EDZ] patterns. For standard *aiurează* 'rants, talks nonsense–3SG’, speakers throughout Muntenia și Dobrogea[^33] vary between the standard form with [EDZ] (*aiurează*) and a form with [ESC] (*aiurește*). See the map below:

![Map of verb distribution](image)

**Figure 4.2:** Adaptation of *Atlasul lingvistic român pe regiuni, muntenia și dobrogea* (1996), harta 100, charting Std. DR *aiurește* 'rant–3SG.PRES.IND.ACT’

[^32]: *Atlasul lingvistic român pe regiuni: Muntenia și Dobrogea* (1996), (harta 23)

[^33]: Similar variation is seen in the corresponding map for this word in the Oltenian dialect atlas (harta 134), which covers the area directly to the west of Muntenia și Dobrogea.
As is clear in the map above, different speakers have categorized this verb differently. For some, it followed the a-stem/[+EDZ] pattern, while for other it follows the i-stem/[+ESC] pattern. While one may be able to spot vague trends in the map above, e.g., there are more forms with [ESC] in the western half of Muntenia și Dobrogea than in the eastern half, etc., we can not make any generalizations.

4.4.4. Variation in contemporary Daco–Romanian

Following Labov (1994), looking at a contemporary language can give us insight into the processes that were active in earlier stages. The following are examples found on the internet\(^34\) of two recent borrowings from English into Daco–Romanian:

\((1)\) **DOWNLOAD**

a. *Descarca* sau vezi pe internet animeuri?

"Do you download or watch anime on the internet?"

\(^{34}\)www.blue-anime.myforum.ro/~vp31751.html- accessed 15 Nov. 2007

b. …\(\text{când}~\text{vreau}~\text{să}~\text{downloadești}~\text{ceva}…\)

"When I want to download something…"

\(^{34}\)forum.chip.ro/archive/index.php/t-81580.html- accessed 15 Nov. 2007

c. …\(\text{când}~\text{am}~\text{început}~\text{să}~\text{downloadești}~\text{ceva}…\)

"When I started to download…"

\(^{34}\)forum.softpedia.com/lofiversion/index.php/t38280.html- accessed 15 Nov. 2007

\((2)\) **BLOG**

a. …\(\text{bloghezi}~\text{asteptand}~\text{trenul.}\)

"...I'm blogging while waiting for the train"

\(^{34}\)dna29.blogspot.com/2006/05/despre-ce-mai-bloghez-astazi.html-accessed 15 Nov. 2007

b. \(\text{Am}~\text{început}~\text{să}~\text{bloghezi}~\text{astăzi}..\)

"I started to blog...

\(^{34}\)perspectivecrestine.wordpress.com/2007/05/26/- accessed 15 Nov. 2007

---

\(^{34}\) Note that these have not been edited, and appear exactly as found online (i.e., without diacritics)
In both of the above cases, it is clear that speakers have formed verbs using three different techniques. In the examples in (a), speakers have formed a verb from the English words utilizing the a-stem/[+EDZ] pattern. In the examples in (b), speakers have formed verbs following the pattern of the [+ESC] group of the 4th conjugation. Perhaps most interestingly, in the examples in (c), speakers have added the affix –ui–
, which automatically makes the verb an i-stem, specifically of the [+ESC] group. The cases in (a) and (c) are the most common options, though it depends on the verb in question. For example, an internet search for verbs formed from English blog resulted in over 12,000 hits for bloguiesc, over 2,000 hits for blog(h)ez, while just over 300 for blog(h)esc. On the other hand, for download there are nearly 50,000 hits for downloadez, 84 for downloadiesc and only 57 for downloadesc. There is not a consistent option that is chosen as the highest rate for both of these verbs, and the fact that one pattern is dominant for each of these not really significant here; what is important is that speakers of DR have made these connections.

It is clear that different DR speakers are forming new verbs via (at least) three different techniques. What did speakers do in the past? There is no reason to say anything other than they did precisely the same thing. Different speakers followed different patterns. In those cases, there often is still variation, and there probably was

---

35 This affix is a denominative suffix of Slavic origin, e.g., OCS věra ‘belief’ + uj/ova (depending on syllabification) → e.g., věrujo ‘believe–1SG.PRES.’, věrovati ‘believe–INF’. It is still active is Modern Slavic as a verb formation strategy, including forming verbs from modern loanwords, e.g., Eng. to park → Rus. parkuju ‘park–1SG’, parkovati ‘park–INF’. For more on this issue, see Rosetti (1962), Costanzo (2008).
36 These searches were conducted on 10 December 2010 and restricted to websites written in Romanian.
37 There are some additional hits for forms like daunlodez, blaghez, etc., that have been adapted to Daco–Romanian orthography.
more variation in the past that was eliminated over time (especially as DR was standardized).

This entire chapter focused on the presence or absence of these stem extensions within lexical items, and completely assuming that the paradigmatic distribution is regular without exceptions. As a side note, I have found cases on the internet of these stem extensions being present paradigmatic slots where they are not expected given the distribution in Standard DR (as well as the trends across Romance discussed earlier in this chapter). The following examples, after observing at phenomena in other varieties of Romance (particularly French as well as non-standard varieties of any other Romance languages), is not a shock. The following examples show the suffix [ESC] in forms where it would not be expected paradigmatically.

(3) *Daca dorești să adaugi un produs in cosul de cumpăraturi...*
    if want-2PL SA add-2PL a product in basket-DEF of purchases
    'If you want to add a product in the basket of purchases...'  
    www.meloegrano-it.ro/meg/?q=content/cumpar - accessed 16 Apr. 2010

The verb *a dori* 'want-INF' follows, in Standard DR, the *i*-stem/[+ESC] pattern. While 2PL forms do not show the stem extension, this case, *dorești*, does (cf. Std. DR *doriți*).

(4) *Când vorbim la telefon nu le putem vedea starea de spirit...*
    when speak-1PL on telephone not PN can-1PL see-INF state-DEF of mind
    'When we talk on the phone we cannot see the state of mind...'  

Again, *a vorbi* 'speak' is, in Standard DR, an *i*-stem/[+ESC] verb. While the Standard form would be *vorbim*, this example shows a 1PL with the affix, *vorbescim*. While more data is necessary to really see what is going on here, it looks like DR speakers (just like some Catalan, Italian, etc. speakers) are extending this suffix throughout the paradigm. Forms like these are rare, but the fact that a DR speaker (I think we can be fairly confident that something written in DR on the internet is written by a native/fluent DR
speaker) extended\textsuperscript{38} this suffix is relevant, and should be investigated in further detail elsewhere.

4.5. **The Other Balkan Romance Languages**

As discussed at the beginning of this chapter, Daco–Romanian is not alone in the Balkan family of Romance languages. One of the more significant factors in the differentiation of this group is that their developments are marked by the different languages with which they have come into contact. As discussed earlier, the $i$-stem/\([+\text{ESC}]\) pattern of the Daco–Romanian is highly populated with verbs of foreign origin, or verbs formed from nouns (regardless of origin). In the other three Balkan languages, the corresponding patterns are followed by largely different inventories of verbs (due to being in contact with different languages), but the same processes appear to have been active (perhaps to different extents, however).

4.5.1. **Aromanian**

As in Daco–Romanian, in the traditional view of the AR conjugalional system (e.g., Vrabie 2000, Ianachieschi–Vlahu 2001) the traditional 1\textsuperscript{st} ($a$-stems) and 4\textsuperscript{th} ($i$-stems) conjugations are each broken into two subclasses: one composed of verbs conjugated the stem extensions [ESC]/[EDZ], and the other without it.

The $i$-stem/\([+\text{ESC}]\) group of Aromanian differs greatly from that of Daco–Romanian, mainly due to the fact that since the split of "Common Romanian", speakers of Aromanian have come into contact with speakers of languages that Daco–Romanian speakers have not (or at least to a different extent). Due to their being situated primarily in Greek and Albanian–speaking territory, Aromanian shows a great number of verbs borrowed from these languages. Turkish has also played an important role in the

\textsuperscript{38} I have also found putative examples of this suffix being extended to the infinitive as well as the imperfect.
formation of the Aromanian lexicon, given that the Aromanian–speaking territory was
under Ottoman rule for centuries. The following are just a few of the examples of words
of these origins being borrowed into Aromanian (Papahagi 1963):

<table>
<thead>
<tr>
<th>Grk.–1SG</th>
<th>tsiknizo 'fret'</th>
<th>AR–1SG</th>
<th>țicnusescu 'fret'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vrondó 'rumble'</td>
<td></td>
<td>vrundusescu 'rumble'</td>
</tr>
<tr>
<td>Alb.–1SG</td>
<td>dëgjoj 'hear'</td>
<td></td>
<td>dăguescu 'hear'</td>
</tr>
<tr>
<td></td>
<td>qëndroj 'stay'</td>
<td></td>
<td>kindrescu 'halt'</td>
</tr>
<tr>
<td>Trk.–1SG</td>
<td>bastirmak 'press'</td>
<td>băstruescu 'scald'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kazanmak 'earn'</td>
<td></td>
<td>căzănsescu 'earn'</td>
</tr>
</tbody>
</table>

Table 4.14: Balkan borrowings in Aromanian

While most of the Balkan borrowings entered the [+ESC] subclass of the 4th conjugation,
there appear to be some exceptions (as in Daco–Romanian).

There are also cases of variation in Aromanian. Most of this variation seems to
be of denominative verbs of Latin origin. Several examples of “double forms” are given
for a–stems that vary between [+EDZ] and [−EDZ], e.g., AR arug’edzū [+EDZ] ~ arog [−
EDZ] ‘engage’, aumbredzū [+EDZ] ~ aumbru [−EDZ] ‘shade’ (note that in Std. DR, this
verb is i–stem/ [+ESC], e.g., DR umbresc). There is variation between [+ESC] and [−ESC]
within i–stems as well: e.g., AR amput [−ESC] ~ amputsăscu [+ESC] ‘stink’, pat [−ESC]–
pătsăscu [+ESC] ‘suffer’ (Capidan 1932). Vrabie (2000) and Papahagi (1963) also give
numerous examples of variation in their dictionaries, both between [+/−] classes of the
i–stems and a–stems, as well as between i–stems and a–stems.
There are also some interesting examples of differences between Daco–Romanian and Aromanian when it comes to neologisms. While the available Aromanian resources do not mention more recent neologisms (such as those discussed in Daco–Romanian in 4.4.3., above), the dictionaries give a few examples of earlier neologisms. For example, the verb 'to photograph' in standard Daco–Romanian is a fotografia, which follows the a-stem/ [+EDZ] pattern (1SG fotografiez). In Aromanian, on the other hand (at least according to some sources), the 1SG is futugrafsescu, which clearly entered follows the i-stem/ [+ESC] pattern. The same trend is seen in the verb for 'telegraph' with 1SG telegrafiez in Daco–Romanian but tiliyrafsescu in Aromanian.

While these are pseudo–neologisms, these examples show evidence for some somewhat recent borrowings entering Aromanian following the i-stem/ [+ESC] pattern. The immediate source of these words in the two languages is different, as both are normally claimed to be from French in Daco–Romanian, while immediate Greek origin is probable for Aromanian (e.g., /y/, a voiced velar fricative, Greek <γ>, appears in

---

<table>
<thead>
<tr>
<th>a-stem/[–EDZ] vs. a-stem/[+EDZ]</th>
<th>i-stem/[–ESC] vs. i-stem/[+ESC]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndreptu ~ ndreptedz 'correct'</td>
<td>mug ~ mudzescu 'bellow'</td>
</tr>
<tr>
<td>turbu ~ turbedz 'go mad'</td>
<td>l'uftu ~ l'uftuescu 'fight'</td>
</tr>
<tr>
<td>a-stem/[+EDZ] vs. i-stem/[+ESC]</td>
<td>tuşedz ~ tuşescu 'cough'</td>
</tr>
<tr>
<td>aumbredz ~ aumbrescu 'cast shadow'</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.15: Variation in Aromanian (Vrabie 2000)

---

39There variation in the forms given in the few available Aromanian dictionaries, and it is important to remember that speakers do not necessarily do what dictionaries say. I have found a case of fotografiadzâ in an issue of Bana Armâneasca. Additionally, in the Macedonian/Aromanian dictionaries published by the Unia tî culturî–a armînjlor dit Machidunii (Cuvata 2006, 2009), this verb is listed the Aromanian volume as fotografescu, but in the Macedonian volume the translation of Mk. fotografira is given as fotografseadzâ. In my fieldwork on dialects of Aromanian spoken in and around Bitola, Macedonia, when prompted with the verb 'photograph', speakers always gave a compound, e.g., fac cadru 'make a picture'. When asked about forms like fotografiadzâ and fotografescu, they found them acceptable but said that they would not say these words when speaking Aromanian.

121
‘telegraph’). It is not particularly clear what happens with more recent borrowings into Aromanian. To try to investigate this, I searched through several issues of the Aromanian magazine Bana Armănească40 to look for cases of neologisms. There many cases of neologisms, most of which appear to be $a$–stem/ [+EDZ], e.g., $\text{developeadzâ}$ ‘develop–3SG’, $\text{ridiculizeadzâ}$ ‘ridicule–3SG’, $\text{transmeteadzâ}$ ‘transmit–3SG’, $\text{continueadzâ}$ ‘continue–3SG’, etc. However, there was also a form $\text{cuntinueshti}$ ‘continue–3SG’ which appears to be [+ESC]. Perhaps the $a$–stem/ [+EDZ] pattern is more productive in Aromanian in actual use than it appears to be by just looking at dictionaries.

When creating a standard where there was none before, speakers may look toward a closely related language that is already standardized. These verbs could have been incorporated from Daco–Romanian, and then “Aromanianized” (changing DR $\dot{z}$ > AR $dz$, raising unstressed mid vowels, etc.). While influence from French may have promoted the productivity of the [+EDZ] class of the 1$^{\text{st}}$ conjugation in Daco–Romanian, it is definitely possible that Daco–Romanian influence is promoting the productivity of the [+EDZ] class of Aromanian.

4.5.2. Megleno–Romanian

As in Aromanian and Daco–Romanian, the traditional views of the Megleno–Romanian conjugational class system are marked by the split of the traditional 1$^{\text{st}}$ and 4$^{\text{th}}$ conjugations into two groups depending on the presence or absence of the [EDZ] or [ESC] suffixes, respectively. (Capidan 1932, Atanasov 1990). As in Aromanian, the $i$–stem/ [+ESC] class of Megleno–Romanian is highly populated with borrowings from the languages with which its speakers have come into contact. In this case, most are of Slavic origin, while there is also a substantial amount of words of Greek origin as well.

---

40 Available online at http://www.geocities.com/armaneasca/.
While all four Balkan Romance languages share certain Slavic words that were incorporated before the split of Common Romanian, Megleno–Romanians have maintained a high level of contact with speakers of Slavic languages (chiefly Bulgarian/Macedonian) throughout the centuries following the Slavic invasions, and the split of Common Romanian to groups north and south of the Danube. These borrowings seem to all have joined following the \i-\text{stem}/[+ESC] pattern:

<table>
<thead>
<tr>
<th>Mk.–3SG</th>
<th>MR–1SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>piše</td>
<td>pisăies</td>
</tr>
<tr>
<td>kači</td>
<td>discâtšos</td>
</tr>
<tr>
<td>tura</td>
<td>tures</td>
</tr>
<tr>
<td>mava</td>
<td>măies</td>
</tr>
</tbody>
</table>

Table 4.16: East South Slavic borrowings into MR

In his recent work on modern Megleno–Romanian, Atanasov (1990) writes:

Today in Megleno–Romanian, the only type of productive verbs are those represented by the verbs of the 4\text{th} conjugation. The enrichment of this conjugation is made directly from Greek or Macedonian, or via their intervention, (Atanasov 1990:210)\textsuperscript{41}

There are several examples of words of French words in Megleno–Romanian (mainly neologisms), but, at least according to Atanasov 1990, they passed through either Macedonian or Greek before entering Megleno–Romanian. This also makes the assumption it appears that the \a-\text{stem}/[+EDZ] is not productive, which is markedly different from Daco–Romanian (where it appears to be the most productive) and Aromanian (where there appears to be some productivity, albeit less than in Daco–Romanian).

\textsuperscript{41}"Aujourd'hui, en mégleno-roumain, les seuls types de verbes productifs son ceux représentés par les verbs de la 4\text{e} conjugaison. L'enrichissement de cette conjugaison se fait directement du grec ou du macédonien, ou par leur entremise"
The scarce supply of Megleno–Romanian sources does not provide much in terms of variation within Megleno–Romanian conjugational classes. However, comparing Megleno–Romanian verbs with their cognates in Aromanian and Daco–Romanian (where cognates are available) shows some variation in conjugational class affiliation of certain verbs. For example, for the verb ‘divide’ (from Lat. *IMPARTIŌ*), MR has a [+ESC] form *ampărtso*, while AR (*ampartu*) and DR (*împărt*) have the same verb, but categorized as [-ESC]. For the verb meaning ‘suffer’ (from Lat. *PATIŌ*), MR and DR agree on having a [+ESC] form (*patsos, pățesc*[^42]⁵⁴, respectively), compared to AR, which has a [-ESC] form, *pat*.

As for the *a*–stem/ [+EDZ] class in MR, it appears to be quite small when compared to that of Daco–Romanian. As discussed earlier, it can be proposed that French borrowings that entered DR and followed the *a*–stem/ [+EDZ] pattern perhaps “revitalized” the productivity of the class, leading it to be the predominantly productive class in the modern language (although there is variation). In Aromanian and Megleno–Romanian, on the other hand, French did not serve as a language of prestige. Any borrowings of French origin likely passed through the dominant language of the area. Thus, there was nothing to ‘revitalize’ the [+EDZ] class of the 1⁰ conjugation, which led to its being apparently ‘less’ productive in Aromanian (at least in terms of neologisms) and not productive in Megleno–Romanian (at least according to Atanasov 1990). For more on this, see Costanzo (2010).

### 4.5.3. Istro–Romanian

Istro–Romanian, spoken by less than one thousand speakers in the Istrian Peninsula, is commonly thought to be connected somewhat closely to Daco–Romanian, and it is presumed that the split of Istro–Romanian and Daco–Romanian is rather recent, or at

[^42]: Daco–Romanian also has a non–standard variant *pat* [-ESC]. This verb shows quite a bit of variation across Romance. Its outcome(s) in Sicilian are discussed in Chapter 6.

124
least much more recent than the split of Common Romanian into South–Danubian and North–Danubian subgroups. Although currently politically part of Croatia, the Istrian Peninsula has been shifted between Yugoslav/Croatian and Venetian/Italian control (as recently as 1947) for centuries. Due to this situation, Croatian, Italian and Slovene (albeit to a lesser extent) have strongly influenced Istro–Romanian both lexically and grammatically, e.g., the maintenance of the infinitive in Istro–Romanian could be attributed to Italian or Croatian influence (Kovačec 1971).

Istro–Romanian is not particularly thoroughly described, and many of the available sources consist mainly of texts which, while valuable, make the facts surrounding the language tougher to access. What is clear is that there is a i–stem/[–ESC] pattern, into which most borrowings fall (including at least one case of a traditional 1st conjugation Italian verb, It. cambiare ‘change–INF’ \(\rightarrow\) IR cambeai, Sârbu (1992). There also appear to be different subgroups (within the 4th conjugation) into which different borrowings from Croatian enter, which appears to be partially dependent on their phonology, but also their aspect. Istro–Romanian (like Megleno–Romanian to a lesser extent, and like Daco–Romanian and Aromanian, to a far lesser extent) has developed a somewhat complex aspectual system under Slavic influence, see Klepikova 1960, Hurren 1969, Kovačec 1971, Mişan 1973 for details).

The a–stem/[+EDZ] pattern present in the other Balkan Romance languages appears not to be present in Istro–Romanian, as far as I can tell. If it is (if the texts and grammars written are not comprehensive, which may likely be the case), then it appears not to be productive. For example, the verb a lucra, which in the other three Balkan Romance languages shows the [EDZ] affix, does not show it in Istro–Romanian, or at least in the speech of the speakers interviewed in the studies. For example, a text in Sârbu (1993) contains the phrase … pórcu če škoda lucra… “The pig that makes damage”. In this phrase, lucra is a 3SG.PRES.IND.ACT form, while we would expect
lucrează in Daco–Romanian, Aromanian and Megleno–Romanian. Other examples of verbs that show the [EDZ] suffix in Daco–Romanian or the other Balkan Romance languages do not show it in the Istro–Romanian texts in Sârbu (1993) and Cantemir (1959) and no mention of it is given in Kovačec’s Descrierea istoromânei actuale.

4.6. Summary and Conclusions

4.6.1. Superclasses?

As discussed earlier, Juilland & Edwards (1971) describe the Romanian conjugational class system with respect to the similarities that certain classes share with other classes. While only five conjugations are proposed (again, based on the infinitival endings –a, –ea, –e, –i, –î), one of the goals of this work is to try to figure out what these classes have in common, e.g., phonological alternations, morphological alternations, etc. The authors look at certain features, and see which of the traditional five conjugations “agree” with one another. For different features, there are different agreements among the classes. For example, for the characteristic of having a verbal extension, the traditional conjugations in –a, –i, and –î agree, as they all have some members that show a verbal extension in certain paradigmatic slots. The authors look at a number of other features, which show agreement across some conjugational classes, and then propose that this motivates the proposal of “superclasses” (classes that cover more than one of the traditional conjugational classes).

While I agree with Juilland and Edwards that the similarity that certain verbs have with other verbs is based on a number of features, and that more than the theme vowel of a verb needs to be evaluated in its categorization, I strongly disagree that this warrants the proposal of “superclasses”. What I do believe these connections show is that there is organization of these patterns that is based on a number of features (theme
vowel, presence of suffixation, etc.), that are better viewed as a “constellation” (Joseph & Janda 1986).

4.6.2. What does this show about conjugational classes?

One might say that the variation between $i$-stem/$[+ESC]$ verbs and $i$-stem/$[-ESC]$ verbs shows that these should be subclasses of a single conjugational class. However, perhaps the variation is just dependent on changing a verb to a different pattern with which it shares a certain characteristic, rather than necessarily warranting subbranching from a single node.

This is evident in the $i$-stem/$[+ESC] \sim a$-stem/$[+EDZ]$ variation seen in DR and AR. There is substantial variation here, and for many verbs both forms exist. No Romance linguist would propose these two patterns branching from a single conjugational class. The reason that there is so much variation between these classes is perhaps that they do share a characteristic, namely that they both are conjugated with a stem extension.

So, if the reasoning for classifying $i$-stem $[-ESC]$ and $i$-stem/$[+ESC]$ together as ‘$i$-stems’ is due to the fact that they share certain characteristics (i.e., they are both $i$-stems) and that there are frequent shifts between the classes, similar reasoning applies for classifying the $i$-stem/$[+ESC]$ and $a$-stem/$[+EDZ]$ together, because they share certain characteristics (they both have a stem extension in the same paradigmatic distribution), and there are frequent shifts between classes.

This is not possible with the traditional model of conjugational classes in Romance. Once there is the split into the traditional four (or five) classes, then there is the implication that there is nothing significant that some members of one class share with some members of another class. This implication is incorrect. What is needed is a model of conjugational classes that allows for similarity based on several factors.
situation is remedied by the model proposed in Chapter 1. I consider these traditional ‘subclasses’ to be independent classes, i.e., there is no internal-class branching. However, these classes are organized according to similarity. The different classes of Balkan Romance have certain similarities. One of these is the theme vowel, but it is by no means the only aspect of these verbs used in categorization by speakers. Nor should it be thought of as primary.

4.6.3. Summary

This chapter discussed the development of verbal patterns in Balkan Romance that are typically referred to as subclasses in the literature. This data clearly shows us that categorization of Balkan Romance verbs is not entirely, or even primarily, dependant on the ‘theme vowel’. The maintenance of the “theme vowel” as the main categorization feature is a holdover from Latin (where it was not incredibly successful either) which biases the description of its daughters. The patterns of variation and change in Balkan Romance show us that there are other features that must be significant in categorization of Balkan Romance verbs. This cannot be described using the traditional system, and the model provided in Chapter 1 proves advantageous.

In addition, the historical evidence shows that the accumulation of verbs following certain patterns is a process that was done in a verb-by-verb fashion. There is also evidence for this in the geographical distribution of these forms, as well as modern techniques used in adapting loanwords. Speakers are not making broad generalizations in any of these cases. Rather, the generalizations speakers made and continue to make are “local” in nature.
CHAPTER 5:
VARIABLE EXTENSION OF VELAR AUGMENTS IN CATALAN

5.0. INTRODUCTION

A frequently studied issue in Romance historical morphology has to do with the emergence of a velar consonant in several verbs in certain paradigmatic slots in several Romance languages, e.g., Lat. TENEÔ ‘hold–1SG.PRES.IND.ACT’ → Sp./It. tengo. In these cases, the emergence of the velar consonant is not etymological, that is, its development cannot have been the result of a regular sound change. While there is a handful of verbs that show this unetymological velar (henceforth referred to as the “velar augment”) in Spanish and Italian, this increases to a great extent in Standard Catalan, and even moreso in non-standard Catalan dialects.

This chapter discusses how this development in Catalan verbal morphology has serious consequences for the traditional notion of “conjugational class” in Catalan, which, as in Latin and Romance in general, is far too dependent on the notion of “theme vowel”, particularly in defining what is or what is not a “class”. The data utilized here come from a variety of sources and allow for an observation of both the earlier stages and the modern situation in Catalan.

The number of patterns involving the velar is large, in terms of patterning both within paradigms and within the lexicon. In all, the distribution of these velar consonants, both within the paradigm as within the lexicon show that there are no broad generalizations to be found in this data. Alternatively, this data demonstrate that, both diachronically and synchronically, local generalizations are at play.
5.1. PRELIMINARIES

5.1.1. The velar augment

Unetymological velar consonants (with a variety of realizations) appear in a number of verb forms in Spanish, Catalan and Italian, e.g., it is found in the 1SG present indicative active, as well as throughout the subjunctive paradigm (to differing extents) of the Romance reflexes of the Latin verb forms VENÌÔ ‘come–1SG.PRES.IND’ and VENIAM ‘come–1SG.PRES.SUBJ’.

<table>
<thead>
<tr>
<th>Lat.</th>
<th>Sp.</th>
<th>Cat.</th>
<th>It.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VENÌÔ</td>
<td>vengo</td>
<td>vinc</td>
<td>vengo</td>
</tr>
<tr>
<td>VENIAM</td>
<td>venga</td>
<td>vingui</td>
<td>venga</td>
</tr>
</tbody>
</table>

Table 5.1: Outcomes of with Lat. VENÌÔ/VENIAM with velar augment

The appearance of this velar is unexpected (and thus, can be considered “unetymological”) as the regular outcome of a Latin prevocalic N+front vowel sequence in Romance is a palatal nasal [ɲ]. In Portuguese, much of Rhaeto–Romance, and Sicilian (as well as other Southern Italian dialects), this is indeed the result in their respective reflexes of Lat. VENÌÔ/VENIAM:

1 In this sense, "etymological" here means "caused by regular sound change" (lautgesetzlich)
2 This phenomenon is not restricted to these languages, as it is also found widely in all stages of Occitan (though the velar in the 1SG has been leveled out in modern Occitan). This perhaps ‘links’ the Catalan and Italian phenomena – allowing this be seen as something present in a certain contiguous area of Romance). It is also seen in Old French, though it has been leveled out in the modern language.
3 In Std. Sp. and Cat., it is found throughout the present subjunctive paradigm. In Std. It., it is only found in the singular and the 3PL, e.g., Sp. vengamos, Cat. vinguem, cf. It. veniamo ‘come–1PL.PRES.SUBJ.’ Also, Std. It. shows the velar in the 3PL.PRES.IND. while Cat. and Sp. do not, e.g., It. vengono, cf. Sp. vienen, Cat. vénen.
The velar result cannot be the result of regular sound change in these languages, as in nouns, the result is a palatal nasal across all of Romance. For example, the reflexes of Lat. ARANA/EUS ‘spider’ all have a palatal nasal (e.g., Ptg. aranha, Sp. araña, Cat. aranya, Fr. araignée, It. raggio, etc.), while hypothetical forms such as *aranja are unattested.

The appearance of the velar augment in the verbal, but not nominal, forms show that a morphological (i.e., analogical) solution is called for, rather than a phonological solution.

The velar augment is present in some Romance languages, while absent in others. Within the languages in which it is present, there is variation in the extent to which it is present. The examples in Table 5.1., above, are cases of a single verb that shows the velar in the 1SG.PRES.IND/SUBJ forms across all of the languages in question. However, these are exceptional cases, as different languages and different dialects show rather different patterns of velar insertion. Spanish and Italian, the peripheral members of this group, show the velar in a limited number of verbs while Catalan and Occitan, spoken in the intervening area, expand this group substantially.

The expansion reaches its greatest extent in Romance in Catalan, as the majority of the traditional Catalan 2nd conjugation\(^5\) verbs in the standard language have a unetymological velar in their 1SG.PRES.IND form (as well as in the subjunctive, etc.). Moreover, in some nonstandard varieties of Catalan, the velar has been extended to an even greater extent. In Spanish and Italian, the verbs that show the velar are usually

---

\(^4\) Most Sicilian dialects do not have a present subjunctive (the present indicative is used in these cases), though if it did, we would expect vegna. Port. \(nh\) Rmsh., It. \(gn\) represent a palatal nasal [\(n\)] (cf. Sp. \(\tilde{n}\), Cat. \(\tilde{ny}\), etc.)

\(^5\) Under traditional views, Catalan 2nd conjugation is the outcome of the Latin 2nd and 3rd conjugations (essentially, verbs in \(-e\)). For more, see Section 5.3.1.
considered to be “irregular”. However, the vast number of verbs that show it in Catalan makes deeming such verbs “irregular” more difficult, as saying these are irregular is basically stating that most verbs of the traditional Catalan 2\textsuperscript{nd} conjugation are irregular\textsuperscript{6}. Thus, the velar augment has, in a way, wreaked havoc on the traditional conjugational class system, and accounts of the traditional Catalan 2\textsuperscript{nd} conjugation (if it is maintained that something called the 2\textsuperscript{nd} conjugation exists) must be broken down into several subclasses to account for the different patterns (e.g., as seen in Badia i Margarit 1994). While in Standard Catalan the velar augment almost always seen in the 2\textsuperscript{nd} conjugation, there are a few cases of its presence in other traditional conjugations. In some non-standard dialects, there has been further extension, thus resulting in even more fragmentation of the traditional conjugational class system.

5.1.2. Focus of previous studies

Much of the early work of the literature on this phenomenon (e.g., the early historical grammars, as well as Malkiel 1974, Lenfest 1978) deals predominantly with unetymological velar augments in Spanish and (albeit to a lesser extent) Italian. Despite the fact that they have a much wider range of verbs showing the velar augment, Catalan and Occitan are rarely mentioned in these studies, and when given more than a brief mention, the situation in non-standard dialects is normally ignored.

As well as neglecting Catalan in general, most previous studies of velar inserts in Romance have generally focused most of their attention on explaining the ultimate origin of this velar consonant. Some studies have dealt with the early stages of analogical spread, but again, this discussion is mostly concentrated on Spanish. While a brief overview of the origin of the velar and its earliest stages of spread is summarized

\textsuperscript{6} For example, in Catalan linguist Pompeu Fabra’s (1926) book of verb patterns, he gives a sample paradigm of the 2\textsuperscript{nd} conjugation, and then says that very few verbs actually follow this model. He then directs the reader to the section on irregular verbs in the back of the book.
briefly below, they are not discussed in a great amount of detail. These are compelling
tions that definitely deserve attention, but my focus here is to describe the
alogical spread of this velar in non-standard Catalan dialects.

There are more recent studies that deal directly with this phenomenon in
Catalan, e.g., Wheeler (1993), Viaplana (2005), Lloret (2009), Querol (2009), as well as
many others. While many are concerned with the emergence of the velar in certain verb
forms, others are more concerned with the synchronic distribution in certain dialects.
While these all are valuable works, most of these studies still maintain the traditional
view of Catalan verbal classes. The focus here is that the emergence and extension of
the velar augment has fragmented the traditional conjugational class system, and
demonstrates the need for a model that does not focus primarily on the theme vowel of
a verb in categorization.

5.2. EARLY DEVELOPMENTS

5.2.1. Origins and initial stages of spread

It has been claimed that the emergence of this velar is analogical to other verbs that
already (etymologically) had the velar in their PRES.IND. form, such as Lat. DICō ‘say’,
FACIō7 ‘make’, as well as the group of verbs in –NGō, such as PINGō ‘paint’ TANGō ‘touch’,
TINGō8 ‘dye’ (Badia i Margarit 1951). This innovation appears to have begun in verbs
such as TENEō and VENIō (and likely REMANEō ‘remain–1SG.PRES.IND.ACT’) given textual
evidence and the fact that they appear consistently with a velar across these languages.
It was then likely extended to other verbs with stem-final nasals, such as Lat. PONō
‘place’ in this stage as well, as all of these languages that do have a reflex of Lat. PÔNÎRE

7 Although Latin FACIō ‘make–1SG.PRES.IND.ACT’ gives Cat. faig [faʧ], which is normally considered
“irregular”, but does not show the velar augment in any form, e.g., feci ’1SG.PRES.SUBJ’, fiu ’1SG.PERF.IND’, fet ‘past participle’.
8 Although the modern Catalan reflexes of the Latin verbs in –NGō do not show the velar augment
(at least in the standard language), e.g., Lat. TINGō gives Cat. tenyir ‘dye–INF’, 1SG tenyeixo.
show the velar in the 1SG PRES.IND. form (Sp./It. *pongo*)\(^9\). The next step involved the extension of the environment to verbs with other stem-final sonorants, such as Lat. SALLŌ ‘jump’ and VALLŌ ‘be strong’, which give Sp./It. saltō and saltō, respectively. We know that this came in a later stage because there are attested forms of these verbs without the velar (e.g., O.Sp. valo). In Standard Spanish and Italian this is approximately where the extension ended (Malkiel 1974, Fondow 2010).

1SG PRES.IND. forms showing variation between palatal (etymological) and velar (unetymological) outcome are attested for some of the languages in question in some of these verbs. For example, from Lat. TENEŌ there is attested variation between tegno and teno in Old Italian and between tenh and ten in Old Occitan. This alternation shows that there was a period of time before the velar option (and in the case of Occitan and French, before a different solution) was settled upon. Perhaps this alternation also occurred in early stages of the languages that settled on the velar option (e.g., Spanish, although there is no textual evidence of such a development) or the palatal option (Portuguese, Ladin, etc.) as well.

Looking at Modern Occitan 1SG.PRES.IND forms is somewhat misleading as analogical change has eliminated the final velars present in earlier stages of the language (e.g., Old Occ. tenc \(\rightarrow\) Mod. Occ. teni). 1SG.PRES.IND forms with unetymological velars are attested in older stages of Occitan (Pellegrini 1954, Skårup 1997) and similarly to Catalan, the number of members in the group is much more expansive than in either Italian or Spanish. In addition, as discussed above, the velar is maintained in the Occitan subjunctive paradigm (e.g., tenga ‘have–1SG.PRES.SUBJ’)\(^10\). In Catalan, the extension of the velar augment into 1SG 1SG.PRES.IND.ACT forms has gone

---

\(^9\) This could also be from a Proto–Romance form *PONEO*, and thus could have been part of the original initial change, along with TENEŌ, VENIŌ, etc.

\(^10\) A similar situation is seen in many Catalan dialects. There appears to have been some paradigmatic pressure to level out the 1SG velar in the indicative, but as the velar was already widespread in the subjunctive, there was less pressure (and consequently, no leveling).
substantially further than in any other Romance language. In Standard Catalan, the velar
has been extended to the majority of verbs descending from the Latin 2nd (–ERE) and 3rd
(–ERE) conjugations, while this is not the case in Spanish or Italian.

<table>
<thead>
<tr>
<th></th>
<th>Sp.</th>
<th>Cat.</th>
<th>It.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCRIBÓ</td>
<td>escribo</td>
<td>escri</td>
<td>scrivo</td>
</tr>
<tr>
<td>CREDÓ</td>
<td>creo</td>
<td>crec</td>
<td>credo</td>
</tr>
<tr>
<td>POSSUM</td>
<td>puedo</td>
<td>puc</td>
<td>posso</td>
</tr>
</tbody>
</table>

Table 5.3: Comparing three common verbs between Cat. and Sp./It.

5.2.2. Catalan and Occitan

A question that can be raised concerns why Catalan and Occitan have developed in this
way to the exclusion of the rest of Romance. The answer is likely found in the perfect
tense. Meyer–Lübke (1923) proposes that the Latin perfect ending in –UI could have
regularly developed into forms with velar consonants in Occitan and Catalan, e.g., Lat.
POTUIT ‘can–3SG.PERC.IND.ACT’ > Occ. poguèt, Cat. pugué. This does not explain the
entire picture of this phenomenon in Romance, as Italian and Spanish show the velar
augment in some (albeit a limited number of) verbs, and there is no evidence that the
Latin perfect ending in –UI developed into something containing a velar consonant in
those languages (e.g., Lat. POTUIT > Sp. pudo, It. poté(tte). While the presence of the
velar consonant in the perfect paradigm cannot be not solely responsible, it could be
seen as a potential reinforcement in Catalan and Occitan. The influences discussed
above that led to the formation of the limited group of verbs showing the
unetymological velar in Spanish and Italian is also in play here, but its reinforcement by
the perfect forms is what likely led to the massive enlargement of this class in Catalan
and Occitan.

Even if we take the above brief account for granted, the issues having to do with
the spread to its modern–day distribution in modern Catalan dialects are still mainly
unresolved. Surely, there are several factors involved in this expansion (and the
differences observed in different dialects). The focus here is not on the origin of the
velar augment, nor the stages of its spread (for opinions on its spread in Catalan, see
Badia i Margarit 1951, Moll 1952). The concern here has to do with the distribution of
the velar in various dialects of Catalan, both within the lexicon and within the paradigm.

5.3. **Velar Augments and the Catalan Conjugational Class System**

5.3.1. **Traditional view(s) of Catalan conjugational classes**

As briefly discussed in Chapter 3, the evolution of the Latin infinitival types (resulted as
follows in Catalan (the numbers indicate the traditional conjugations of Catalan).

<table>
<thead>
<tr>
<th>Latin</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ĀRE</td>
<td>-ĒRE</td>
<td>-ERE</td>
<td>-ĪRE</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Catalan</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>-âr</td>
<td>-èr</td>
<td>-er/-re</td>
<td>-ír</td>
</tr>
</tbody>
</table>

Table 5.4: Evolution of Cat. infinitives/conj. classes

Catalan maintains the distinction in the infinitive between traditional Latin 2\textsuperscript{nd} and 3\textsuperscript{rd}
conjugations. Latin -ĒRE results in Catalan -èr, while Latin -ERE results in Catalan -re or
-er (both unstressed!). The difference between -er and -re is due to the phonological
structure of the root (Badía Margarit 1951). However, all verbs that have an ending
involving -e (regardless of if it is stressed or its position) are considered to make up a
large “2\textsuperscript{nd} conjugation”. If just paying attention to the infinitive, Badia i Margarit (1951)
describes the situation as follows:

1\textsuperscript{st} conjugation: Verbs ending in -ar
2\textsuperscript{nd} conjugation: Verbs that do not end in -ar (1\textsuperscript{st} conj.), nor in -ir (3\textsuperscript{rd} conj.)
3\textsuperscript{rd} conjugation: Verbs ending in -ir (but dir\textsuperscript{11} is 2\textsuperscript{nd} conjugation)

\textsuperscript{11} Cat. dir ‘say-INF’ is considered to part of the class with no theme vowel (like dur, fer, etc.) that
also form part of the broad 2\textsuperscript{nd} conjugation.

136
The 1st and 3rd conjugations are fairly straightforward. However, Badia i Margarit (1951) describes the 2nd conjugation as comprising pretty much anything that does not fit squarely in the 1st and 3rd conjugations. He then goes on to describe the subclasses of the 2nd conjugation (again, solely based on the infinitive):

2a: Verbs with unstressed ending –re
2b: Verbs with unstressed ending –er
2c: Verbs with stressed ending –er
2d: Verbs ending in –r

Thus, there are really six types of infinitives in Catalan. However, the situation (clearly) is much more complicated when going beyond the infinitive. Later, when discussing the entire conjugational system, Badia i Margarit (1994) still maintains the three conjugational class system, and again breaks down his 2nd conjugation, albeit in a different way than showed above. He leaves his classes IIc and IIId alone, as they both have few members. Then he combines IIa and IIb, but then breaks this class down into five groups (some of which actually have to be split up again).

Group A: Verbs in –re and in –er that are not included in any of the four other groups, (a) in –re (e.g., rebre ‘receive’, perdre ‘lose’, (b) in –er (e.g., tèmer ‘fear’, esprémer ‘squeeze’, etc.).

Group B: Verbs that have have an infinitive ending in –re preceded by a d, which are those that end in –ldre (e.g., molder ‘grind’) and those in –ndre (e.g., ofendre ‘offer’, dependre ‘depend’, etc.).

Group C: Verbs in –re preceded by u (with etymological labial; also identifiable because the u corresponds to Spanish b or v), (e.g., beure ‘drink’, escriure ‘write’).

Group D: Verbs in re preceded by u (with origin other than labial; equally identifiable because the u does not correspond to Spanish b or v), (e.g., caure ‘fall’, plaure ‘cry’, etc.).
Group E: Verbs in –er preceded by a palatal (ix): we can distinguish a first subgroup (Ea: conèixer ‘know’, parèixer ‘seem’) and a second subgroup (Eb: nàixer ‘be born’, créixer ‘grow’, etc.).
(Badia i Margarit 1994:573–4)\(^{12}\)

Badia i Margarit then goes on to admit that this classification is not exhaustive. If we take this information, though, and add to it the fact that the traditional Catalan 3\(^{rd}\) conjugation (verbs in –ir) is traditionally divided into two classes, one showing a reflex of Latin [ESC] in certain forms (e.g., llegir ‘read–INF’, llegixo ‘read–1SG.PRES.IND’) and the other in which it is absent (e.g., dormir ‘sleep–INF’, dormo ‘sleep–1SG.PRES.IND’), the following picture of Catalan conjugation starts to emerge:

![Figure 5.1: A traditional view of Catalan conjugational classes](image)

Of course, this does not mention any “irregular” verbs, or groups of verbs that have alternations that, while phonologically predictable (they are usually governed by stress),

\(^{12}\) “Grup A: Verbs en –re i en –er no compresos en cap dels quatre grups restants (a) en –re (e.g. rebre, perdre, etc.), (b) en –er (e.g. tèmer, esprémer, etc.); Grup B: Verbs l’infinitiu dels quals acaba en –re precedent d’una d’eufònica, que són els que fan –ldre (e.g. moldre) i els que fan –ndre (e.g. ofendre, dependre, etc.); Grup C: Verb en –re precedent de u (amb labial etimològica; també identificables perquè la u hi correspon a b o v castellanes), (e.g. beure, escriure); Grup D: Verbs en –re precedent de u (d’origen altre que labial; igualment identifiables perquè la u no hi correspon a b o v castellanes) (e.g. caure, plaure, etc.); Grup E Verbs en –er precedent de palatal (ix); podem distingir-hi un primer subgrup (Ea: conèixer, parèixer) i un segon subgrup (Eb: nàixer, créixer, etc.)."
the specific verbs of which they are a property are not predictable. Accounts similar to this are the norm for Catalan – two fairly cohesive conjugations (1st and 3rd) and then a highly fractured 2nd conjugation.

5.3.2. Conjugations and velar augments
As discussed above, in Standard Catalan, the velar augment is mainly found in the traditional 2nd conjugation. However, the velar augment is not restricted to the Catalan 2nd conjugation. It appears in two common verbs of the 3rd conjugation, tenir ‘have–INF’ (1SG tinc) and venir ‘come–INF’ (1SG vinc) as well as one verb of the 1st conjugation, estar ‘be–INF’ (1SG estic). Thus, if one believes that there is something like a 2nd conjugation in Catalan, it cannot be seen as being a rigid boundary for the existence of the velar augment. Given these facts, one could potentially propose a “velar class” of verbs that is not dependent on the theme vowel, as nearly all proposed conjunctional systems of Romance languages are.

However, any branching diagram, such as that shown in Figure 5.1., ignores the fact there are verbs across all three traditional conjugations that share this feature. This diagram is entirely dependent on the theme vowel of the verb for the major classification, and any other features are considered secondarily. As is shown in the following sections, there is quite a bit of variation across Catalan dialects (as well as within areas themselves) in that the velar augment appears in the traditional 1st and 3rd Catalan conjugations. However, there appears not to be “conjugation–wide” extension, which shows that speakers are not making extensions based solely on thematic vowels. Thus, it appears that there are other factors at play in the extension of the velar augment, and thus, the general categorization of Catalan verbs as well.

Catalan, even more so than Balkan Romance, as discussed in Chapter 4, has undergone changes that make it difficult to fit it under the traditional notion of
Romance conjugational class. We could definitely split the verbs up into three classes, but one of these three classes, at least according to Badia i Margarit, has to be fractured into at least seven additional classes (a classification that he himself says is not comprehensive). The traditional Catalan 2nd conjugation does show some consistent features, but the defining features of this conjugation are fairly arbitrary and based in tradition, rather than the facts of the language.

Following a brief description of the data sources utilized in this project, the specific facts of the velar augment in Catalan and the dialects thereof are presented in the remainder of this chapter. First, the extent of the velar is discussed in terms of lexical range, showing that a vast number of verbs that do not show the velar augment in Standard Catalan do so in dialects. Second, the distribution of the augment within paradigms is discussed, showing that some dialects diverge from the standard language as to the paradigmatic distribution of the velar.

5.3. SOURCES OF DIALECTAL DATA

5.3.1. Atlas lingüístic de Catalunya (ALC)
The first, and most limited for the purposes of this study, source of data is Antoni Griera's Atlas lingüístic de Catalunya (henceforth ALC), originally published in 1923 (2nd ed. 1966). As is the case with many early dialect atlases, it is heavily based on J. Gilliéron and E. Edmont's Atlas linguistique de la France (1912). Griera’s atlas (as well as the other principal sources of data used here) covers the entire Catalan–speaking area (“els Països Catalans”) which extends far beyond the borders of the modern Spanish autonomous community of Catalonia, into other areas of Spain (the Comunitat Valenciana, the Balearic islands, western Aragon), as well as Southeastern France (Catalunya Nord), the principality of Andorra, and Alghero (Cat. L’Alguer), a city in Northwestern Sardinia.
Griera selected 101 points throughout the Països Catalans and interviewed one (and in a few cases, two) speaker from each between 1912 and 1922. The medieval religious centers and the centers of the Catalan-speaking comarques were chosen, as they, in the words of Griera, “conserves the character of the dialect”. One could dispute the accuracy of Griera's findings as some of the aspects of the study (as well as those of any study following under the category of “dialect geography”) can be seen as problematic. Even though a more consistent sample of speakers could have been interviewed, the data collected and the subsequent dialect maps are still a valuable source of information showing how dialectal Catalan was spoken in the early 20th century.

While the project covers hundreds of individual lexical items, the number of conjugated verbal forms found in the atlas is quite limited. In some cases, particularly with verbs considered irregular, e.g., ‘be’, a wide number of paradigmatic forms are listed. For other verbs, this is not the case, and only a few forms are charted. The ALC uses an older transcription system. In order to make this consistent, I have adapted these transcriptions into IPA format given the descriptions of the symbols used in Volume 1 of the Atlas.

5.3.2. La flexió verbal en els dialectes catalans (FVDC)

Perea (1999) is an electronic version of the data collected by Alcover and Moll for their volume on verbal inflection in Catalan dialects (Alcover & Moll 1929). This data was collected between 1906 and 1921 and covers the entire Catalan-speaking area. This source covers a wide range of verbs (approximately 70 in all), though some are covered

---

13 For example, the ideal informants were generally “NORMS” (non-mobile, older, rural males). In addition, some of the informants were probably far too aware of the standard language/dialectal differences, e.g., Griera himself, as well as Pompeu Fabra.
in much detail than others (there is data from a few of the verbs from only a few locations).

One issue with this data is that the individual speaker from which each form was taken is not identified in the FVDC. This is especially problematic in cases where a number of variant forms are given in one location. While this shows us that in many cases there is variation, it is difficult to get much of a perspective on the verbal systems of the individual speaker. For example, for data point 21, the following variants are listed for the 1SG.PRES.IND of the verbs collir ‘collect-INF’ and cosir ‘sew-INF’

<table>
<thead>
<tr>
<th>Std. Cat. INF</th>
<th>collir</th>
<th>cosir</th>
</tr>
</thead>
<tbody>
<tr>
<td>Std. Cat. 1SG.PRES.IND.</td>
<td>cullo</td>
<td>cuso</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>variant 1SG.PRES.IND. forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>FVDC, point 21</td>
</tr>
<tr>
<td>cullo</td>
</tr>
<tr>
<td>culloc</td>
</tr>
<tr>
<td>culli</td>
</tr>
<tr>
<td>cullic</td>
</tr>
<tr>
<td>colleixi</td>
</tr>
<tr>
<td>colleixic</td>
</tr>
<tr>
<td>colleixo</td>
</tr>
<tr>
<td>colleixoc</td>
</tr>
<tr>
<td>cuso</td>
</tr>
<tr>
<td>cuseixo</td>
</tr>
<tr>
<td>cuseixi</td>
</tr>
<tr>
<td>cusi</td>
</tr>
</tbody>
</table>

Table 5.5: Variant 1SG.PRES.IND. forms of collir, cosir; FVDC point 21

Clearly, there quite a bit of variation within this small geographical area. However, it is not possible to investigate further the behavior of particular speakers, as the specific variant forms are not assigned to a specific speaker (something that is remedied in the other source of data discussed below). One could definitely make reasonable assumptions (e.g., that some/many of the speakers that have 1SG culli also have 1SG cusi), but these can not be proven whatsoever. Also, one might be tempted to say that the four forms that collir has that cosir does not are just the four forms they share with the addition of the velar augment to each (e.g., culli, culli-; culleixo, culleixo-; cusi, etc.), and this might show that cosir is a verb that just does not have the velar augment.
Perhaps this is true for these speakers, but a look at the surrounding areas shows many variants of cosir with a velar augment.

5.3.3. **Corpus Oral Dialectal (COD)**

Viaplana, et al. (2007) is an electronic corpus that covers various aspects of Catalan dialects, including a section on verbal conjugation. The data is considerably more modern than the other two sources, having been collected between 1994 and 1996. Several verbs are covered in detail (though not to the same extent as the FVDC, as the COD is not solely focused on verbal morphology).

An added advantage of the COD is that several (usually 3–4) speakers from each of the selected areas are represented, which allows to get a clearer perspective of the variation within each area. In addition, each verbal form is marked with a specific speaker in a specific area, which allows one to get a better view of the individual conjugational systems of each speaker represented. This will be the primary source of data used in discussing the paradigmatic distribution of the velar, as it provides a wealth of forms throughout the paradigm, and allows us to see variation within single speakers, as well as variation between several speakers from the same area.

5.3.4. **Other sources of dialectal data**

There are several other sources of data used in this chapter, though to a far lesser extent. While many Catalan grammars focus solely on the standard dialect of the language, some also discuss other dialects of Catalan. Griera (1949) is essentially a review of what he found in the dialect atlas project. For each major area, he gives the main conjugational patterns, but glosses over many of the details. In addition, Griera completed a few dialect atlases focusing on smaller areas within the Països Catalans, giving the opportunity to look at variation in more detail in these areas than in the ALC.
These include Andorra (1960) and the Val d’Aran\(^{14}\) (1967). There are also several Valencian grammars that occasionally stress the dialectal differences between Valencian and standard Catalan (often touching on verb morphology, and specifically the velar augment), e.g., Salvador (1978).

There are several sources that specifically provide conjugational patterns in Catalan, though nearly all of these focus on the standard language. For example, Fabra (1926) which gives the main patterns according to the traditional classification of Catalan conjugational classes (which is essentially the same as the traditional pattern of Romance). However, this becomes an issue, especially in what he refers to the “second conjugation” (again, essentially the group of verbs that do not have an infinitive in \(-ar\) or \(-ir\)), where very few verbs actually follow the model he gives to illustrate it. There are some similar sources (though of a limited number) that cover different dialects, most notably Renat i Ferris (1933), listing the conjugational patterns in Valencian.

A final source of data of information on Catalan verb conjugation comes from various documents advocating the use of a certain variety as the standard. There is a series of publications by the Institut d’Estudis Catalans (or the Institut d’Estudis Valencians) that propose an oral standard of the Catalan language, e.g., IEC (1999) covers morphology and in the discussion advises against the use of certain forms, including many which show the velar augment where the standard does not, e.g., for the verb *morir* ‘die–INF’, there is a note saying that “the first person singular form *morc* is not recommended” (cf. Std.Cat. *moro*).

Such documents are like a modern Appendix Probi, a 3\(^{rd}\) century AD document listing Vulgar Latin forms and their “grammatical”, classical equivalents. The fact that these forms are prescribed against gives us evidence of their existence.

\(^{14}\) While the Val d’Aran is politically part of Catalunya, the dialects spoken there (Aranese) are normally not considered to be Catalan dialects. Rather, Aranese is usually considered a variety of Gascon (a variety of Occitan), spoken across the border in Southern France.
5.4. DIALECTAL VARIATION IN THE EXTENT OF THE SPREAD OF VELAR AUGMENTS

5.4.1. Focus

This section is simply discussing the “lexical appearance” of the velar, that is, a verb is considered “velar” if it shows it in the 1SG.PRES.IND. and “not velar” if it does not. This is following a simple generalization (later shown as false) that if the velar augment appears in the 1SG.PRES.IND of a verb, it also appears throughout the subjunctive paradigm, the past participle, the simple perfect, etc. In this section the sole focus is the 1SG.PRES.IND. While this can be seen as irresponsible, the reason behind this has to do with the availability of data, particularly in the ALC. As discussed above, the ALC presents few conjugated verb forms, and of the conjugated verb forms, many are in the 1SG. In addition, the nature of the data in the FVDC does not really allow us to make any assumptions of the paradigmatic distribution of the velar, as it does not show indications of individual speaker behavior. A more in-depth discussion of the paradigmatic distribution of the velar (using data from the COD) is provided below in section 5.5. (and an even more in-depth view can be seen in Querol 2009).

In addition, here the focus is only be on 1SG.PRES.IND. variants showing the velar augment. Given the history of these forms, there is going to be quite a bit of variation. The Latin 1SG.PRES.IND ending -ō was regularly lost in word-final position in Catalan. In most dialects, a new ending was reintroduced, but the form of it varies (usually –o/–u in Central Catalan, usually –e/–i Northern and Southern varieties) (Badía Margarit 1951). For example, the following reflexes of Lat. cantō are found in ALC, map no. 389:
All that we are really concerned about are the velar forms that exist in some areas. In this case, [kántuk] (showing the velar augment) occurs in a small area in the northeast versus other forms (those without the velar augment) that occur elsewhere in the Països Catalans. The variation found in the other forms (in the final vowel, as in [kánt–u], [kánt–o], [kánt–e] and [kánt–i], or its absence, as in [kánt]) is interesting, though not relevant at this time.

5.4.2. (Near) Cross–dialectal consistency

A preliminary trend that can be seen with the data in the ALC and FVDC is a verbs has the velar augment in the standard language (which, only by looking at the ALC, seems the most conservative when it comes to the spread of the velar augment), that verb usually appears with the velar augment throughout the Catalan dialects (again, see Querol 2009 for a more in-depth discussion). For example, in the ALC, dialectal variants of Std. estic ‘be–1SG’ shows up consistently with the velar (though there is some variation to the quality/presence of the initial vowel, as well as a final vowel present in some dialects). The FVDC also shows this trend, though there are a few points in northern dialects that have forms without the velar (e.g., estau, estai). Most of the exceptions to this trend are in northern dialects where it looks like the velar is absent in some verbs that show it in Std. Catalan. However, different dialects show different patterns. The table below shows the 1SG.PRES.IND. of three verbs (all of which...
have the velar augment in Std. Cat.) found in the first three locations in the FVDC (forms with velars are *italicized*):

<table>
<thead>
<tr>
<th>INF</th>
<th>caure ‘fall’</th>
<th>coure ‘cook’</th>
<th>creure ‘think’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.PRES.IND. Std. Cat.</td>
<td>caic</td>
<td>coc</td>
<td>crec</td>
</tr>
<tr>
<td>1SG.PRES.IND. FVDC Loc. 1</td>
<td>caui</td>
<td>coui</td>
<td>crec</td>
</tr>
<tr>
<td>1SG.PRES.IND. FVDC Loc. 2</td>
<td>caui</td>
<td>coc</td>
<td>crec</td>
</tr>
<tr>
<td>1SG.PRES.IND. FVDC Loc. 3</td>
<td>caui</td>
<td>coui</td>
<td>creui</td>
</tr>
</tbody>
</table>

Table 5.7: 1SG.PRES.IND of *caure/coure/creure* – FVDC loc. 1,2,3

The three locations show different patterns with respect to the appearance of the velar in these three verbs. In addition, there is variation within some of the locations covered in the FVDC with respect to these three verbs:

<table>
<thead>
<tr>
<th>INF</th>
<th>caure ‘fall’</th>
<th>coure ‘cook’</th>
<th>creure ‘think’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG.PRES.IND. Std. Cat.</td>
<td>caic</td>
<td>coc</td>
<td>crec</td>
</tr>
<tr>
<td>1SG.PRES.IND. FVDC Loc. 17</td>
<td>caui</td>
<td>coc</td>
<td>crec</td>
</tr>
<tr>
<td>1SG.PRES.IND. FVDC Loc. 20</td>
<td>caui</td>
<td>caic</td>
<td>cau</td>
</tr>
</tbody>
</table>

Table 5.8: 1SG.PRES.IND of *caure/coure/creure* – FVDC loc. 17,20

The verbs that appear (for the most part) cross-dialectally with the velar augment are nearly all *e*-stems (of the traditional Catalan 2\textsuperscript{nd} conjugation). Again, there are a few exceptions. As discussed earlier, there here are two *i*-stems that show up in Standard Catalan showing the velar: *tinc* ‘have–1SG.PRES.IND’ (cf. INF *tenir*), *vinc* ‘come–1SG.PRES.IND’ (cf. INF *venir*) (though some consider them to be *e*-stems, despite the infinitival –*i*), and one *a*-stem: *estic* ‘be–1SG.PRES.IND’, (cf. INF *estar*). However, this is
not the case with all dialects of Catalan, as some dialects have extended the velar into other non-\(e\)-stem verbs (i.e., \(a\)-stems and \(i\)-stems) without regard to conjugational class boundaries.

5.4.3. Extension into \(a\)-stem verbs

The \(a\)-stem verbs (the traditional Latin/Romance ‘1\(^{st}\) conjugation’) constitute by far the largest traditional group of verbs across Catalan and is widely considered the most stable of the Latin/Romance conjugations. As discussed in Chapter 3, there is a pan-Romance reputation of stability in the 1\(^{st}\) conjugation, as members of this group are not typically shifted between conjugations, stem allomorphy within paradigms is often avoided\(^{15}\), and it is not as frequently affected by analogical change in general.

However, as in Balkan Romance and in other (mainly somewhat peripheral) Romance languages, evidence does show that this conjugation does not deserve such a reputation\(^{16}\). In Catalan, the velar has been extended into the 1SG PRES.IND. of some verbs of the 1\(^{st}\) conjugation in some (most notably northeastern) dialects. For example, the ALC has two maps\(^ {17}\) for 1SG forms of \(a\)-stem verbs, both of which have variants showing unetymological velar consonants in the northeast.

In standard Catalan, the 1SG.PRES.IND of the verb *cantar* is *canto* [kán.to]. However, in some areas in northern Catalunya, the form [kántuk] is found. The presence of the velar is undoubtedly analogical to the other 1SG.PRES.IND forms with

---

\(^{15}\) For example, in the traditional Italian 1\(^{st}\) conjugation (i.e., verbs with infinitives in –are), there is no stem allomorphy within the present paradigm when the stem is followed by \(-i\), e.g., *cercare* ‘look for’, 1SG *cerco*, 2SG *cerchi* (without palatalization) vs. traditional 2\(^{nd}\) conjugation *vincere* ‘win’, 1SG *vinco*, 2SG *vinci* (with palatalization).

\(^{16}\) Analogical change is not as foreign to the 1\(^{st}\) conjugation as is often thought. This claim probably has resulted in only observing the standard dialects of the major Romance languages. However, observation of the lesser-studied Romance languages (particularly Sardinian and the various Rhaeto–Romance and Balkan Romance languages) show that this conjugation is indeed quite frequently subject to analogy; examples provided in Chapter 3.

\(^{17}\) *canto* ‘sing’-1SG.PRES.IND.: Map 389, *dono* ‘give’-1SG.PRES.IND.: Map 660
the velar augment in these dialects. The extent to which unetymological velars are present in other 1\textsuperscript{st} conjugation 1SG PRES.IND. for these speakers is not particularly clear. Griera (1949) implies that not all first conjugation 1SG.PRES.IND. forms have the velar in the dialects where [kántuk] is found. In a summary of the “Eastern Catalan” dialect, he explains that first conjugation verbs with stem vowel [a] have the velar (e.g., \textit{cantar}), while those with the stem vowel [o] do not. Thus, when listing the possible 1SG forms for the verb \textit{trobar} (to find), a form with a final velar (i.e., *trobuk) is not listed.

However, Griera’s own data calls this claim into question, as the verb \textit{donar} ‘give’ is listed in his atlas. This verb has a stem vowel [o] and its 1SG.PRES.IND. forms in the northern dialects indeed do show unetymological velars. However, in this case, there are two different outcomes involving the accretion of the velar augment.

<table>
<thead>
<tr>
<th>Std. Cat.</th>
<th>dono</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variant 1:</td>
<td></td>
</tr>
<tr>
<td>Catalunya Nord</td>
<td>[dónuk]</td>
</tr>
<tr>
<td>Variant 2:</td>
<td></td>
</tr>
<tr>
<td>Girona</td>
<td>[dóŋk]</td>
</tr>
</tbody>
</table>

\textbf{Table 5.9: Velar variants of Std. Cat. dono, ALC 660}

An interesting point is that while the areas in which the 1SG of ‘sing’ shows a velar and the areas where the 1SG of ‘give’ shows a velar overlap to a certain extent, but are not identical. Taking the velar variants ([dónuk] and [dóŋk]) together, all of the speakers interviewed for the ALC that have a velar variant of ‘sing’ also have one of ‘give’. However, the reverse is not true, as several speakers have a velar in ‘give’ but do not have one in ‘sing’ (e.g., Speaker 102 has [kánti] but [dónuk]). This is interesting because it shows that some speakers may classify these verbs together, but some definitely do not. Also interesting is the fact that, if we take the velar variants of Std. \textit{dono} separately, then the overlap is mainly between speakers with [dóŋk] and speakers
with [kántuk]. Only two speakers that have [dónuk] also have [kántuk] (the rest have [kánti]).

While these data come only from two dialect maps, they show that the extension of the velar augment into a-stem verbs has not been a wide-spread change in these varietites. Speakers have extended the velar to these two verbs, but have not done so in precisely the same way. I assume that some of this variation is phonological in nature, e.g., the augment can be directly added to [don–] without a problem, while adding it directly to [kant–] would violate some phonotactic constraint, thus warranting epenthesis, see Costanzo 2006). However, there is still variation in the way that this velar is added (in the case of donar), as well as its presence in general (some speakers have it in their variant of dono but not in canto).

As for the other sources, their focus strays away from a-stem verbs, probably due their reputation for ‘regularity’. Since this broad group is normally viewed as regular, most researchers of Catalan (or Romance in general) tend to push it to the side and focus more on the classes with a more notorious reputation for being irregular. This leads there to be much less data being collected concerning verbs in these classes. All three of the sources used here discuss the forms of cantar, and beyond that, usually only the traditionally described irregular verbs that happen to be a-stems are given much attention (e.g., estar ‘be’, anar ‘go’).

5.4.4. Extension into i-stem verbs
According to the ALC, and as far as I can tell, the FVDC and COD, a-stem verbs that show the velar augment in the 1SG.PRES.IND. are geographically restricted to Northeastern Catalan dialects. Along with the extension of the velar augment into a-stem verbs, it has also been extended into i-stem verbs (traditional 3rd Catalan conjugation) in some dialects. However, unlike the extension into a-stems,
1SG.PRES.IND. i-stems showing the velar are found in a wide range of Catalan dialects, most notably most of the political Comunitat Valenciana, as well as several of the areas of Northern Catalunya where velar-augmented forms of a-stems are found. The ALC contains a few maps of 1SG.PRES.IND. forms of i-stem verbs, e.g., one charting dialect variants of Std. cuso, the 1SG.PRES.IND. of cosir ‘sew’. See the simplified\(^\text{18}\) map below:

As shown in the map/table above, there are three different forms showing the velar in three different areas of the Països Catalans. Variants 1 and 2 involve the velar apparently directly added to the root kus- (with a reintroduced final vowel in Variant 2).

\(^{18}\) The maps presented here have been simplified to solely focus on the forms showing the velar augment. Circles have been drawn around forms showing the same pattern of accretion of the velar augment. I am not making the assumption that all speakers from these areas have these forms. As seen in the FVDC, there is often a tremendous amount of dialectal variation within these surveyed areas. These circles only indicate where all of Griera’s interviewees had the same response.
Variant 3, on the other hand, looks like it involves the addition of the velar after the reintroduction of the final vowel.

Of the set of Griera’s maps showing the variability of 1SG forms, the case that stands out as showing the most variation is that of dormo. The map shows four different analogical outcomes involving the addition of a non-etiymological velar, as well as other analogical possibilities. ¹⁹

<table>
<thead>
<tr>
<th>Standard Catalan</th>
<th>[dɔ́ɾmu]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variant 1: Valencia</td>
<td>[dɔ́ɾk]</td>
</tr>
<tr>
<td>Variant 2: SE Tarragona</td>
<td>[dɔ́ɾmiyo]</td>
</tr>
<tr>
<td>Variant 3: Catalunya Nord</td>
<td>[dɔ́ɾmik]</td>
</tr>
<tr>
<td>Variant 4: Girona</td>
<td>[dɔ́ɾmuk]</td>
</tr>
</tbody>
</table>

Figure 5.3: Std. dormo, adapted from ALC, Map 665

Comparing these two maps (along with comparing them in terms of what is happening in Northeastern Catalunya with the maps of canto/dono) brings out some interesting points. Starting with the south, we see that much of Valencia shows the velar augment in the 1SG of both verbs ([kúsk] and [dɔ́ɾk] respectively). There is a small area in

¹⁹ These mainly involve the verb having undergone a shift to follow to the i-stem/[+ESC]. While dormir usually does not join this group in Standard Catalan (or in Italian or DR, for that matter), it does in some Catalan dialects.
Tarragona where it looks like it was extended, along with the introduction of a final vowel, giving [kúzgo] and [dórmīyo], respectively. However, the extent of the areas differ, as there are speakers that have [kúzgo] but not [dórmīyo]. Similar situations are seen in the north, where there are two different velar–augmented variants of *dormo*, but only one of *kuso* (which shows a much more restricted distribution than the velar forms of *dormo*).

The examples of 1SG.PRES.IND. forms of *i*-stem verbs in the ALC are limited to the two above, but additional sources show that additional *i*-stem verbs exhibiting the velar augment. Renat i Ferris (1933)²⁰ (from approximately the same era as Griera’s atlas) lists a number of *i*-stems that have the velar augment in Valencian, e.g. òbric ‘open–1SG.PRES.IND’, cf. Std. Cat. obro; òmplic ‘fill–1SG.PRES.IND’, cf. Std. Cat. omplo.

The other main sources of data described above show substantial of variation in 1SG.PRES.IND forms with regard to the velar augment as well. All verbs in the FVDC that are *i*-stems have at least several 1SG.PRES.IND forms showing the velar augment. However, there is a great deal of variation. For example, in the ALC, Std. *cuso* has forms with the velar augment in Northern Catalunya, but they are all [kúzuk]. However FVDC also shows forms like *cusic* in these areas, often existing in the same area forms like *cusuc*. Comparing the ALC and the FVDC makes the picture look a bit more complicated, as the FVDC looks at everything in much more detail. When taken together, a rather complicated picture emerges that shows that the extent of the velar augment definitely varies from location to location and speaker to speaker, and broad generalizations about the behavior of speakers throughout the Països Catalans, across speakers in a specific area within the Països Catalans, and even within single speakers,

---
²⁰ Renat i Ferris claims that the extent of the extension in the *i*-stems has to do with the final consonant of the stem, and that stems with final [–m] are not subject to velar extension. He then claims that *dormir* is an exception, as it shows stem–final [–m] and shows the velar throughout Valencia.
are not to be found in this data. There definitely are patterns, but their scope is rather limited, mainly covering a few forms.

5.4.5. Consistency

The table below summarizes the results with regards to the presence of the velar augment in the four 1SG forms of Catalan verbs found in the ALC and discussed above. The shaded areas indicate forms that show the unetymological velar stop in the 1SG.PRES.IND form in question.

<table>
<thead>
<tr>
<th>Std. Cat.</th>
<th>canto</th>
<th>dono</th>
<th>cuso</th>
<th>dormo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valencia</td>
<td>[kánte]</td>
<td>[dóne]</td>
<td>[kús[k]</td>
<td>[dórk]</td>
</tr>
<tr>
<td>Tarragona</td>
<td>[kánk]</td>
<td>[dóno]</td>
<td>[kúzgo]</td>
<td>[dórmuy]</td>
</tr>
<tr>
<td>Girona</td>
<td>[kántuk]</td>
<td>[dónk]</td>
<td>[kúzik]</td>
<td>[dórmuk]</td>
</tr>
<tr>
<td>Catalunya Nord</td>
<td>[kánti]</td>
<td>[dónuk]</td>
<td>[kúzi]</td>
<td>[dórmik]</td>
</tr>
</tbody>
</table>

Table 5.10: Summary of four 1SG forms in ALC

This is the extent of the relevant data in the ALC. Again, as discussed above, there is also a great amount of data in the FVDC that emphasizes that the preliminary findings reported here are quite incomplete. Using these four examples as our concrete evidence, it is not possible to form any strong generalizations about the distribution of the velar augment (which verbs are affected as opposed to those that are untouched) as well as the formation of 1SG forms in these dialects. Any generalization made could easily be shown as lacking in the face of additional data.

21 The terms “Valencia”, “Tarragona”, “Girona” and “Catalunya Nord” do not imply that the boundaries of the different forms found are in any way identical (or close to identical) to the political boundaries that divide these areas. In addition, what I call “Valencia”, “Tarragona”, “Girona” and “Catalunya Nord” in this table actually vary from verb to verb, as there are small differences in the extent of the spread of these forms. These names are just used for convenience, and are attempting to try to make some sort of sense of the distribution of these verbal forms.
Despite the fact that looking solely at these four forms will in no way solve the all the riddles of the dialectal variation in extension of velar consonants in Catalan, a comparison can give some insight. For example, the rough area I call “Girona” is the only area discussed in which all four verbs show a velar consonant in their respective 1SG forms. Three of the four ([kántuk], [kúzuk], [dɔ́ɾmuk]) show a consistent pattern that appears to cross conjugational boundaries. All three end show final [-uk], which points to the simple addition of a [-k] to forms that had a final [u] restored.

Given the three forms shown above, it might be predicted that the 1SG.PRES.IND. of donar found in this dialect would be [dónuk]. However, the form found here is [dóŋk], which lacks the expected restored final [-u]. Why is this form different from the other three? The answer cannot be driven by conjugational class, as donar and cantar are both traditionally considered to be part of the a-stem/1st conjugation. Perhaps the [u] only serves to break up a non-permissible consonant cluster. However, this does not explain why the 1SG form for 'sew' here is [kúzuk], as [-sk] is a perfectly acceptable word-final consonant cluster (e.g., visc [bisk] ‘live-1SG’). It appears that the velar was generalized to this verb before the final vowel was restored. Again, these data show that this is a morphologically driven development, rather than a phonologically-driven one.

The divergent forms found here may be explained through postulating two distinct stages of spread of the velar augment. There is some evidence for such a notion. Recall that the 1SG form of estar in this area (as well as throughout Catalunya) is estic, while forms like [astuk] are unattested. It appears possible that one set of forms (the ones that occur consistently cross-dialectally) acquired the velar in a stage

---

22Cat. donar is an a-stem verb. However, there are some irregularities in that there are some additional forms in the paradigm that are actually forms of the verb dar, which has the same meaning. For example, there are alternate 1PL/2PL forms, e.g., donem (from donar) vs. dem (from dar), imperfect forms, e.g., donava (from donar) vs. dava (from dar), past participial forms, e.g., donat (from donar) vs. dat (from dar), among others. Perhaps donar does not fit into this pattern because its “irregular” in some way. I do not believe this, but I am sure that this might be the opinion of some.
prior to the restoration of the final vowel. After this stage, the final vowels were restored to all but the velar–final forms. After the restoration of the final vowels, the velar was added (via analogy to the first group) to the group of verbs that did have their final vowel restored.

Perhaps in this dialect, the 1SG of donar acquired its velar augment with this first group (along with estic, tinc, vinc, etc.), which may account for the fact that it lacks the restored vowel. The obvious question deals with why donar would acquire its velar in the stage in which it was greatly confined to the 2nd and 3rd conjugations. This definitely could have something to do with frequency, as ‘to give’ is rather frequent, just like the other 1st conjugation verb that the velar augment spread to in this stage, estar. However, there is no way to securely know what happened. What we can see, though, is that this development has been local in nature.

Despite the fact that the different verbs show some differences in velar insertion in Girona, all four have the velar, which superficially makes it look like this has been extended across the board. Interestingly, looking at the area described as Catalunya Nord, a different conclusion can be reached. This conclusion is that it looks like the notion of conjugational class acting as a barrier to velar insertion has completely broken down. One of the two a-stem verbs discussed above has the velar, while the other does not. One of the two i-stem verbs discussed above has the velar, while the other does not. No generalization based on conjugational class can be made here, thus demonstrating that this change is analogical in nature and suggesting that the expansion of this class is accomplished on a form–by–form basis, rather than a class–by–class basis.

Not only does this area show inconsistency in terms of conjugational class affected, but it also does not show any consistency in terms of the formation of the velar–augmented forms. The Valencian and Tarragonese forms could be reconciled
rather easily and the Gironese forms could be reconciled with some effort. The forms found in Catalunya Nord are [dɔ́rmik] and [dónuk], both of which appear to have been formed by the accretion of a velar consonant after the restoration of a vowel replacing Latin/Proto-Romance [o]. However, the quality of the vowel in the two verbs is different. The typical restored vowel in Catalunya Nord is [−i] (e.g., the Catalunya Nord reflex of CANTÔ is [kánti]) making the form [dɔ́rmik] not surprising. However, [dónuk] is definitely not expected, as the apparent restored vowel here is [u]. There is not enough data to postulate that this has something to do with the structure of the verb. This is just further evidence that analogy operates in a form–by–form fashion, and even the group of forms that are ultimately affected may be affected in different ways.

5.4.6. Consistency among speakers

In the above sections, I have labeled dialects as “Gironese”, “Valencian”, etc. However, for each of the verbs in question, the limits of these “dialects” change. For example, the map below shows the limits of the areas where the 1SG of dormir is [dɔ́rmuk] and the 1SG of cantar is [kántuk] (this area was earlier referred to as “Girona”). While these areas overlap a great deal, they are not identical, as some speakers that have [kántuk] lack [dɔ́rmuk] and vice–versa.
There is variability in single speakers as well. For example, in the ALC, one speaker responded as having both [kántuk] and [kánti] along with both [kúzuk] and [kúzi]. A respondent from a different point had both [dőrmuk] and [dőrmu].

Any (albeit premature) generalization that may be made must be constrained due to the findings above. The generalizations seen in analogical change are highly "localized" (Joseph & Janda 1988) in terms of both dialect space and lexical range. Broad, sweeping generalizations are not to be found in this data.

5.5. VARIATION IN THE PARADIGMATIC DISTRIBUTION OF THE VELAR AUGMENT

5.5.1. Focus

The preceding section treated this issue as if the paradigmatic distribution of the velar augment were consistent across dialects and across verbs. That is, verbs were discussed in a binary fashion, i.e., either they had the velar or they lacked the velar based on the 1SG.PRES.IND form. The previous section is valuable in that it does discuss the variation in the presence of the velar and the havoc that this wreaks on the traditional notion of conjugational class in Catalan. However, this is a simplification, as there are paradigmatically different patterns in different dialects, as well as differences
between speakers of the same dialect (even speakers from the same town). There is even variation within individual speakers, i.e., cases in which one speaker has two verbs that show different paradigmatic distribution of the velar. The situation described in the previous section showed that the spread of the velar within the lexicon and its present synchronic status is quite complex. Looking at the details of its distribution in the paradigm brings a whole another level of complexity to the situation.

5.5.2. Patterns in Standard Catalan

In standard Catalan, generalizations can be made as to the appearance of the velar consonant. The following table shows the paradigmatic distribution of the velar consonant in the verb *tenir* ‘have-INF’

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>IND.PRES.</td>
<td>tinc</td>
<td>tens</td>
<td>té</td>
<td>tenim</td>
<td>teniu</td>
<td>tenen</td>
</tr>
<tr>
<td>SUBJ.PRES</td>
<td>tingui</td>
<td>tinguis</td>
<td>tingui</td>
<td>tinguem</td>
<td>tingueu</td>
<td>tinguin</td>
</tr>
<tr>
<td>IND.PERF</td>
<td>tingui</td>
<td>tingueres</td>
<td>tingué</td>
<td>tinguérem</td>
<td>tinguéreu</td>
<td>tingueren</td>
</tr>
<tr>
<td>IND.IMP</td>
<td>tenia</td>
<td>tenies</td>
<td>tenia</td>
<td>teniem</td>
<td>tenieu</td>
<td>tenien</td>
</tr>
<tr>
<td>SUBJ.IMP</td>
<td>tingués</td>
<td>tinguessis</td>
<td>tingués</td>
<td>tinguéssim</td>
<td>tinguéssiu</td>
<td>tinguessin</td>
</tr>
<tr>
<td>FUT</td>
<td>tindré</td>
<td>tindràs</td>
<td>tindrà</td>
<td>tindrem</td>
<td>tindreu</td>
<td>tindran</td>
</tr>
<tr>
<td>COND</td>
<td>tindria</td>
<td>tindries</td>
<td>tindria</td>
<td>tindriem</td>
<td>tindrieu</td>
<td>tindrien</td>
</tr>
<tr>
<td>PART: tingut</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GER: tenint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INF: tenir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 5.11: Conjugated forms of *tenir* ‘have’ in Std. Cat.*

In general, the velar appears in the 1SG.PRES.IND., the entire present and imperfect subjunctive paradigms, the perfect paradigm, as well as the past participle. This is the typical distribution in the standard language. However, even in Standard Catalan, there are some exceptions. First, there are a number of verbs that do not show the velar in the present indicative of subjunctive, but do show it in the perfect, the past subjunctive, and the past participle. Below is the complete set of conjugated forms of the verb *córrer* ‘run’
If we take the Latin –ui→ Cat. g sound change briefly discussed in previous sections as a given, then the paradigm of this verb is more or less what is expected. However, speakers still have to deal with the fact that there are different “velar” patterns, some of which enter the present indicative/subjunctive, and some that do not. There are also several verbs that show the velar in the same pattern as tenir, except that it is absent in the past participle. For example, the verb vendre ‘sell-INF’ shows the velar in the 1SG.PRES.IND. (venc), throughout the present subjunctive (e.g., 1SG venguí), throughout the perfect indicative (e.g., 1SG venguí) and imperfect subjunctive (e.g., 1SG vengués), but not in the participle, venut.

Another pattern that diverges from that of tenir is seen in the paradigm of saber ‘know-INF’, which does not show the velar in the 1SG.PRES.IND. form sé (which is pretty anomalous). The velar is not present in the participle (sabut), the perfect indicative (e.g., 1SG sabí) or the imperfect subjunctive (e.g., 1SG sabés), but it is present in the present subjunctive (e.g., 1SG sàpiga). A similar pattern can be seen in cabre ‘fit-INF’ (e.g., 1SG.PRES.SUBJ. câpiga), though this verb has a more regular-looking 1SG cabo.

Thus, while most Standard Catalan verbs that show the velar follow the same paradigmatic pattern as to its present, there are some other, more restricted patterns. Explanation for several of these variations could be found in the diachrony of the forms, however, since speakers are do not have access to the history of their language as they...
are acquiring it, these would just be different patterns to them. The patterns in the standard language are not completely consistent, but when looking outside of the standard, many different patterns are found. Some of these patterns are somewhat consistent, and some show an astounding amount of variation.

5.5.3. Patterns in other dialects

The following discussion focuses on the presence/absence of the velar in the present indicative and subjunctive paradigms. As discussed above, it is also often present in the perfect, participle, etc. This will not be discussed in further detail here, though a few comments are given at the end of this chapter.

While lacking much information on verb conjugation, the ALC does have a few maps of 1SG.PRES.SUBJ maps corresponding to some of the 1SG.PRES.IND maps discussed above. A comparison of these maps, while giving nothing near a concrete conclusion, does provide some interesting information. Recall that the 1SG.PRES.IND of Std. cantar is [kántuk] in some Northeastern dialects of Catalan, and that in the standard, when there is a velar augment in the 1SG.PRES.IND, it is expected in throughout the present subjunctive paradigm. Interestingly, none of the speakers who have 1SG.PRES.IND [kántuk] have a 1SG.PRES.SUBJ with the velar augment (for all of these speakers, the 1SG.PRES.SUBJ is [kánti]). Perhaps even more interestingly, one speaker does have a 1SG.PRES.SUBJ with a velar, [kánteya], but a corresponding 1SG.PRES.IND [kánto], without the velar.

<table>
<thead>
<tr>
<th>location #</th>
<th>1SG.PRES.IND</th>
<th>1SG.PRES.SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>11, 12, 33, 34, 35, 36, 37</td>
<td>[kántuk]</td>
<td>[kánti]</td>
</tr>
<tr>
<td>5</td>
<td>[kánto]</td>
<td>[kánteya]</td>
</tr>
</tbody>
</table>

Table 5.13: 1SG.IND./1SG.SUBJ. variation re: velar augment; ALC 389/394
A comparison of the maps for Std. *dormo* ‘sleep–1SG.PRES.IND’ and *dormi* ‘sleep–1SG.PRES.SUBJ’, yields more interesting data. Just as with the previous example, the 1SG.PRES.IND. forms with a velar in Northeastern Catalan do not have a corresponding 1SG.PRES.SUBJ. forms with a velar. All of the speakers in Valencia that have the velar in the 1SG.PRES.IND. do have it in the 1SG.PRES.SUBJ. Recall that there was a small area in Tarragona (just 2 speakers in the ALC) that had a 1SG form [dɔ́ɾmiyo]. Both of these speakers have a 1SG.PRES.SUBJ. form with the velar ([dɔ́ɾmiyə]/ [dɔ́ɾmiya]) However, many of the surrounding speakers that did not have the velar in the indicative do have it in the subjunctive. See the chart below:

<table>
<thead>
<tr>
<th>IND and SUBJ w/ velar</th>
<th>location #</th>
<th>1SG.PRES.IND</th>
<th>1SG.PRES.SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>72, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 89</td>
<td>[dɔ́ɾk]</td>
<td>[dɔ́ɾya]</td>
</tr>
<tr>
<td></td>
<td>56, 58</td>
<td>[dɔ́ɾmiyə]</td>
<td>[dɔ́ɾmiyə]</td>
</tr>
<tr>
<td>IND w/ velar, SUBJ w/o velar</td>
<td>11, 12, 13, 14, 33, 34, 35, 36, 37, 53, 54, 102, 103, 104</td>
<td>[dɔ́ɾmik]</td>
<td>[dɔ́ɾmik]</td>
</tr>
<tr>
<td>IND w/o velar, SUBJ w/ velar</td>
<td>5, 6, 7, 21, 22, 40, 59, 69, 70, 71</td>
<td>[dɔ́ɾmo]</td>
<td>[dɔ́ɾmiyə]</td>
</tr>
</tbody>
</table>

Table 5.14: 1SG.IND./1SG.SUBJ. variation re: velar augment; ALC 665/668

There could be several explanations for these developments. As for the speakers that have the velar in the subjunctive but not the indicative, the extension to the subjunctive could have been an earlier step than the extension to the indicative, and the final step of extension to the indicative just never happened. Or potentially, there could have been a velar in the 1SG.PRES.IND that was later eliminated through paradigm leveling (this is what happened in Occitan, and it definitely appears that this could have happened in some dialects of Catalan). Again, we cannot be sure about how this all happened, and it does not concern us here. What does concern us is that there are patterns that diverge
from the norm in Standard Catalan. This is nowhere near the extent of variation in present indicative and present subjunctive paradigms in dialectal Catalan.

For example, the Catalan verb *perdre* ‘lose’, which does not show the velar in Std. Cat. (e.g., 1SG.PRES.IND *perdo*, 1SG.PRES.SUBJ *perdi*), does show the velar to differing extents in some non-standard varieties. The following chart shows the different patterns shown for this verb in the *Corpus Oral Dialectal*. All of these diverge from the standard language in that they do show at least one present indicative or subjunctive form with the velar. While some show the typical pattern shown by verbs that have the velar in the Standard language, others diverge to a much greater extent. In the following chart, the black squares represent a form with the velar augment, while the white squares represent a form without the velar augment.

<table>
<thead>
<tr>
<th></th>
<th>present indicative</th>
<th>present subjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>2S</td>
<td>3S</td>
</tr>
<tr>
<td>1P</td>
<td>2P</td>
<td>3P</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 5.15: Velar patterns in the PRES.IND/SUBJ. paradigms of Std. *perdre* (COD)

Pattern (1) represents the general pattern of velar insertion (at least with respect to the PRES.IND/SUBJ) in Standard Catalan. Pattern (2) has eliminated the velar in the 1SG.PRES.IND, likely due to analogy with the rest of the indicative paradigm (without affecting the subjunctive). Pattern (3) does not show the velar in the 1SG.PRES.IND, but does show it in the 1PL/2PL.PRES.IND. Pattern (4) shows the standard pattern with the addition of it showing up in the 1PL.PRES.IND (but not the 2PL!). This looks kind of odd as the 1PL/2PL often pattern together in Catalan/Romance, and often (at least in the
indicative) sit out of various changes due to being stressed on the ending instead of the stem. However, in this case (along with several of the other patterns shown above, they definitely do not pattern together. Pattern (5) shows the standard pattern along with the extension into the 2SG. Pattern (6) only shows it in the 2PL.PRES.SUBJ. The pattern in (7) shows the standard pattern except it is present in the 2PL.PRES.IND (but not in the 1PL). As for the subjunctive, it is not present in the 3PL. Pattern (8) only shows it in the 1/2PL.PRES.SUBJ, and finally, Pattern (9) shows the standard pattern except it is not present in the 1/2PL.PRES.SUBJ.

The only semi-regularities that can be brought up about this data is that there are no cases (of this verb, in this corpus) that show the velar in the indicative but not the subjunctive (though we have already seen that this does happen in some dialects), and that it looks like many of these patterns are, in some way, derived from the standard pattern. However, does the fact that these look like they are derived from the standard pattern really matter? It definitely does matter for those interested in the specific stages of the paradigmatic spread. However, we, looking at this synchronic state, should focus solely on this synchronic state, and instead of paying attention to how this happened, we should pay attention to how speakers deal with the fact that it is this way. Even though it looks somewhat complex, it is clear that speakers deal with it just fine.

5.5.4. Variation within areas

The data presented in Table 5.15, above, comes from various areas throughout the Països Catalans. However, when looking at areas with a more limited geographical range, there is also quite a bit of variation as to the appearance of the velar. Recall that in the ALC, several speakers gave two different forms when prompted with an inflected form. In addition, there was a lot of variation shown in the FVDC, e.g., eight variants for the 1SG of collir in a single area. The COD also shows quite a bit of variation, and
since individual speakers are identified and many areas are represented by more than one speaker, it is much easier to get a better grasp on variation within a small area. The following, while not necessarily a dramatic example, shows variation across the three speakers interviewed from Alacant with regards to their conjugation of the present indicative and subjunctive paradigms of the verb *sentir* 'hear'.

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>IND</td>
<td>séŋk</td>
<td>sénts</td>
<td>sén</td>
<td>sentém</td>
<td>sentéw</td>
<td>sénten</td>
</tr>
<tr>
<td>SUBJ</td>
<td>sénτa</td>
<td>sénτes</td>
<td>sénτa</td>
<td>sentéτm</td>
<td>sentéτw</td>
<td>sénten</td>
</tr>
</tbody>
</table>

**Table 5.16: PRES.IND./SUBJ. paradigms of *sentir*, 3 Alacant speakers**

Speaker 1 has the velar in the 1SG.PRES.IND and nowhere else. Speaker 2 does not show the velar anywhere in the paradigm. Speaker 3 shows the most interesting distribution, as the stem extension [ESC] is shown in much of the indicative paradigm, while the velar augment is shown throughout the subjunctive paradigm.

### 5.5.5. Variation within speakers

The data from the ALC and the FVDC cannot really tell us that much about the variation within individual speakers (well, we can take it as such, especially with the ALC, but with the FVDC, it’s really touch to know the exact source beyond the location of the forms in question. However, the COD lets us see variation within individual speakers. The following at the present indicative and subjunctive paradigms of two verbs given by a single speaker from Tortosa (Southern Tarragona).

---

23 Even though these forms (at least the 1SG) have a velar consonant, it is not the velar augment, and thus, are not shaded.
Both of these verbs do not show the velar augment in Standard Catalan. This speaker has the velar augment in the subjunctive paradigm in these two verbs. However, the distribution of the forms showing the velar is different. Again, this shows cases of local generalizations in the presence of this velar consonant.

5.6. SUMMARY AND CONCLUSIONS

The data presented in this chapter show that the development of this velar augment in Catalan morphology has had serious consequences for the notion of conjugational class in Catalan. Traditional accounts of Catalan conjugational classes place primary importance on the identity of the theme vowel, normally giving three conjugational classes. However, under these accounts (e.g., Badia i Margarit 1994), as discussed above, the traditional 2nd conjugation is fractured, as various subclasses (and occasionally subclasses of subclasses!) are needed to account for the actual number of patterns that actually exist. Just as in the discussion of Balkan Romance in the previous chapter, these extensive subbranching models lead to making inaccurate predictions. One could ask what the point of a “2nd conjugation” is if its proposal results in the proposal of extensive branching. This does cover a generalization regarding the theme vowel\(^{24}\), but it does not cover anything else, and in the end, it just looks like an attempt to have a model that looks elegant and somewhat Latin-esque on the surface. Note also that all of these traditional accounts are aimed at describing Standard Catalan, while

\(^{24}\)This generalization across the theme vowels is not even particularly clear.
what has occurred in non-standard dialects presents an even more extensively fragmented picture.

One could propose that there is a “velar class” of verbs in Standard Catalan that intersects with the traditional conjugational class system. However, this is problematic for several reasons. If we propose a “velar class”, then how does this fit with the traditional classification by theme vowel? Even though it is not everything, the theme vowel still is important in verb categorization. One could argue that classification by theme vowel is primary, or that classification by the presence/absence of the velar is primary. However, both arguments are insufficient. In addition, not all verbs that show the velar augment show it in precisely the same distribution. This gets even more complex in non-standard dialects, where the same speaker can show different distributions of the velar augment in two different verbs.

A model in which similarity can be shown on different levels is more appropriate for describing this Catalan data. There is some relationship between verbs that show the velar in the same distribution, even if they have different theme vowels. There is potentially also some relationship between verbs that show the velar augment in different distributions, though the relationship is not as close.

The generalizations seen in this data are not widespread. In any dialect, there are going to be verbs that follow the same pattern. However, the vast majority of patterns, if we take the definition of pattern laid out in Chapter 1 seriously, will only be followed by a few verbs at most. These data show cases of very “local” generalizations, something which traditional views of Romance conjugational classes are not prepared to deal with. While the traditional systems of describing these classes may look simpler and more elegant, this is achieved at the expense of the psychological reality of the situation and the facts of the language.
CHAPTER 6:
ON SOME ISSUES IN SICILIAN AND AROMANIAN VERB CONJUGATION

6.0. INTRODUCTION

The following chapter describes several issues pertaining to verb conjugation in specific Sicilian and Aromanian dialects. Fieldwork was conducted to determine how verbs are conjugated in these varieties, with a conscious attempt by the interviewer to not be biased by the traditional views of the classification of verbs into conjugational classes. This chapter does not serve as a complete analysis of verbal inflection in these areas. Rather, after a brief overview of these dialects, a series of specific examples will be given that show that, while speakers definitely do make generalizations, the generalizations are more “local” than broad in scope.

These examples are by no means as “dramatic” as the Catalan and Balkan Romance data presented in the two previous chapters. However, these examples still clearly demonstrate that classifying verbs primarily according to their theme vowel in these dialects is inappropriate and makes incorrect predictions.

6.1. METHODS

6.1.1. Fieldwork

Fieldwork\(^1\) was conducted in Montedoro, Sicily and Bitola, Macedonia during July and August 2009. Approximately three weeks were spent in each area. The dialects studied were a Sicilian dialect in Montedoro and an Aromanian dialect in Bitola.

\(^1\) IRB protocol number 2009E0546
languages were chosen as they often treated as “peripheral” by most Romance linguists, who usually push them to the sidelines or completely ignore them. Within these languages, the specific dialects focused on can also be seen as “peripheral” as they have received limited attention even in studies of these languages. There are specific reasons for the choice of Montedoro and Bitola as the specific areas studied; these are elaborated in detail in Section 6.2. for Sicilian and Section 6.4. for Aromanian.

6.1.2. Speaker Recruitment

The recruitment of speakers was different in each area. While the original intention was to pay informants, every offer of payment was rejected. In Montedoro, the targeted individuals were older people who had lived in the town most (if not all) of their life, and thus, every native Montedorese (nearly the entire population of the town) had potential to be a subject for this study. I have a connection with Montedoro, as my grandparents were born there and immigrated to the United States in the early 20th century. Much of my family remained in Montedoro after my grandparents left, though most, but not all, of their descendants have since left. My family in the United States has maintained fairly close ties with Montedoro and I visited briefly several years earlier. I believed that my connections would give me a jumpstart in finding informants. However, I was met with some reluctance by some potential informants.

I faced a more difficult task in Bitola as the vast majority of the municipality’s population (over 95%) are not speakers of Aromanian. However, while most inhabitants of Bitola are not Aromanian speakers, I found that many of the people I encountered had some Aromanian ancestry, which ended up facilitating my search for informants. One issue is informants’ village of origin was not taken into account in recruiting informants, and thus, the informants interviewed in this study descend originally from different villages within Bitola Municipality. Since Aromanian is mainly used in the home, the
original dialect differences of my informants have been maintained to a certain extent even though they all were born Bitola (except for one who was born in a village outside Bitola but has lived in Bitola most of their life). This led to some complications, as discussed below.

6.1.2. Elicitation

The verbs for which conjugated forms were sought fell into two basic categories. The first category consisted of a list of verbs chosen beforehand that show interesting trends in Aromanian and/or Sicilian (as well as in Romance in general), based on the available literature. These included (1) verbs considered “irregular” in traditional sources, e.g., ‘be’, ‘go’, etc., (2) selected verbs that did not have a consistent root in the present/perfect in Latin or in Romance, e.g., ‘write’, ‘have’, etc., and (3) a representative or two from each traditionally-defined Latin conjugational class, as well as the traditional subclasses within, e.g., e-stem, i-stem/[+ESC], i-stems/[−ESC], etc. This list was designed to elicit some of the inherited vocabulary from Latin, due to the fact that both of these languages (Aromanian to a far greater extent) have undergone great deal of language contact that has resulted in some of the inherited Latin vocabulary being replaced.

The other group of verbs was a list of basic verb meanings with no targeted Aromanian/Sicilian verb in mind when asking them. This group of verbs was by no means fixed and additional verbs were asked on the fly during the interviews. Since the focus of this project was to avoid the bias of the original conjugational class descriptions of these dialects, this was wholly intentional.

For instance, in the Aromanian varieties focused on here, even rather common verbs have been replaced, e.g., ‘speak’ is zbrâscu, which is of Slavic origin (cf. Mk. zboruva ‘speak’), ‘cook’ is mayirpsescu, from Grk. μαγειρεύω, etc.
Before I began to do fieldwork, a series of pictures showing basic verb meanings was prepared to assist in the elicitation process. When the speakers were shown a picture, they were asked to tell me what was happening in the pictures, and then were prompted to conjugate the verb in the present tense and the simple past tense. The pictures were designed principally to be used in Montedoro, as I felt that prompting informants with the verb in Standard Italian could potentially have influenced their response. In Bitola, this was less of an issue, as prompting informants in Macedonian was less likely to influence their Aromanian response, since Macedonian and Aromanian are much less closely related that Sicilian and Italian. However, this could still be an issue as these two languages have been in contact for such a long period of time and Macedonian words are occasionally incorporated into Aromanian (especially in the speech of bilinguals).

As it turns out, using the pictures was not particularly successful. I found that most speakers were much more cooperative when actually prompted with a verb to translate into the language in question. This was definitely the case in Bitola, where, for some reason, speakers were much more reluctant to participate in the interview when I attempted to use the pictures. In these cases I abandoned the pictures, and just prompted the speakers with verbs in Macedonian (or in some cases, English), which they translated into Aromanian and then provided conjugated forms.

The point of this project was such that it was not necessary (or even helpful) to collect data from a large segment of the population. Thus, the number of informants was rather limited, totaling five in each area. Some of the informants were interviewed several times, and others only once, mainly due to their time constraints and whether they were cooperative or not. Each meeting lasted between one and two hours. Despite the length of the meetings, the data collected in each meeting was not very
large, as the informants were often much more interested in telling me stories about their villages rather than conjugating verbs.

6.1.3. Limitations
Collecting this type of data is rather tedious for the interviewee, and some speakers were more cooperative than others. In addition, the data collected does not represent the way that these forms are actually used in conversation. However, the methods utilized here is the only way to collect this type of data. Recording conversation between speakers and then analyzing the verb forms that they used would have been ideal, but this would require hundreds of hours of recording, and still there might not be enough data to come to any sort of conclusion of the patterns that these speakers used in their language.

6.2. Background on Montedoro, Sicilian Dialects, and Sicilian Verb Conjugation

6.2.1. Montedoro, Sicily
Montedoro is a small town (population around 1800 and steadily falling) on the eastern edge of Caltanissetta province in south-central Sicily, approximately equidistant (about 30 kilometers) from the provincial capitals of Agrigento and Caltanissetta, and approximately 100 kilometers southeast of Palermo. Once a booming sulfur mining center in the mid-to-late 19th century, Montedoro has seen its industry die, both due to and resulting in many of the town's inhabitants moving on to work in the Pennsylvania coal mines in the early 20th century.

Montedoro is an ideal location to conduct this type of fieldwork, given its substantial distance from the major urban centers of Sicily (Palermo to the north, Messina, Catania, Siracusa on the east coast), as well as its location with respect to the other towns in the general area. The nearest town, Serradifalco (population approx.
lies approximately eight kilometers away, a distance that seems much longer than it actually is, as the only road to Montedoro winds around a hill. Montedoro is connected to the rest of Sicily by decent roads, especially once one gets down the hill to Serradifalco. A few daily buses connect Montedoro to Caltanissetta along with a daily, though not particularly reliable, bus to Palermo. These factors could have caused the dialect of this area to be less influenced by Standard Italian (unlike some of the more urban dialects). In addition, as stated above, Montedoro has experienced great emigration in the past, but minimal immigration. While over the past century millions of Sicilians have been lured away from their homeland to the brighter economic situations in the Americas or elsewhere in Europe, not many people have chosen to immigrate to central Sicily.

6.2.2. Background on Sicilian dialects

Despite the vast differences between the dialects of Sicilian and some other dialects traditionally classified under the heading of “Italian”, most Romance linguists consider Sicilian dialects to be dialects of Italian. The varieties referred to as “Italian” are best seen as a dialect continuum, ranging from the Alps to Sicily (or even further at both ends, if one considers Romance to be a dialect continuum in itself). The Sicilian dialects have never received any status from the Italian government since the unification of Italy into its more-or-less modern state in the mid 19th century. Recently, particularly with the increased availability of schooling over the past one hundred years, Standard Italian has made its way more and more into Sicily, particularly to the major urban centers, but also inland to the more rural portions of the island.

Besides their documentation in the Italian dialect atlases, the dialects of Sicily have not been given much attention in the Romance linguistics literature. Moreover, much of what has been written about Sicilian dialects is based on the dialects spoken in
the major cities along the coast (Palermo, Catania, Messina, Siracusa, etc.) rather than
the more rural areas in the center of the island. Ducibella (1934) gives a decent
overview of phonological/phonetic differences across different dialects, including some
of the more rural varieties of the center of the island. Much of the research on Sicilian
verb morphology comes from Jaberg & Jud's *Sprach-und Sachatlas Italiens und der
Südschweiz* (1928–40), as well as Rohlfs (1968), which, in part, reports the findings of
the atlas project, along with some historical background on dialectal differences. The
atlas, particularly Volume VIII, of which part is devoted solely to verb inflection, covers a
limited number of verbs – representatives of each conjugation (and subclasses) in
various tenses and the typical “irregular” verbs. While the goals of the dialect atlas were
to obtain a large amount of data that could be compared over all over the surveyed area,
the present project involves drastically minimizing the geographical area studied while
widely increasing the portion of the verbal system under investigation.

In addition, much of the available information about Sicilian dialects is rather
dated. For example, Avolio (1882) has a decent overview but is not really indicative of
the modern situation in Sicily, given that the island has undergone a great deal of
upheaval (socially, and consequently, linguistically) in the past century. Some other
sources (Bonner 2001, Camilleri 2001, etc.) seem to be of a more “popular” nature,
which certainly does not take away any of their worth or usefulness (e.g., some of these
were used to determine a starting set of verbs to be elicited), but they are more
intended either to claim that Sicilian is different enough from Italian to be called a
separate language or to teach the fundamentals of the dialect.

Moreover, most of these recent sources refer to the Sicilian language or the
Sicilian dialect, thus making the assumption that Sicilian is a monolithic entity.
However, there is substantial phonological, lexical, morphological, etc. variation across
the island. My goal here is not to provide a clear, updated picture of variation in verbal
morphology across the island, but rather to provide a clear, detailed view of a single
dialect, particularly one that has likely experienced a limited level of influence from
Standard Italian.

6.2.3. Traditional views of Sicilian conjugational classes

As briefly discussed in Chapter 3, the specific sound changes that Sicilian dialects
underwent led to the traditional 2\textsuperscript{nd}/3\textsuperscript{rd}/4\textsuperscript{th} Latin conjugational class infinitival endings
collapsing, giving –iri throughout (Tagliavini 1963). Several sources, e.g., Camilleri
(2001), thus claim that Sicilian has two conjugational classes based solely upon the
infinitive.

Sicilian verbs are reconciled in two large groups that we call conjugations; the
the first conjugation belong those whose present infinitive ends in –ari; to the
second belong those that end in –iri.

(Camilleri 2001:41)

However, as there are clearly more than two patterns found in Sicilian verb conjugation,
Camilleri then goes on to discuss the many verbs deemed “irregular” in Sicilian – only a
handful from his 1\textsuperscript{st} conjugation but a fairly lengthy list from his 2\textsuperscript{nd} conjugation. These
descriptions are not “bad” by any means, but classifying all Sicilian verbs into two
conjugations entails major subbranching and picking out certain forms out as “irregular”.
Many of these sources are more concerned with teaching the language, so saying that
there are only two classes is probably the smart move, as it will not be overwhelming to
the student.

Other sources (e.g., Pitrè & Wentrup 1995) claim that there are three
conjugations in Sicilian. As discussed in Chapter 3, Latin infinitives in –ÈRE (the

\footnote{3 \textquotedblleft I verbi siciliani sono riuniti in due grandi raggruppamenti che chiamiamo coniugazioni; alla
prima coniugazione appartengono quelli che all’\textit{infinito presente} terminano in \textit{ari}; alla seconda
coniugazione appartengono quelli che terminano in \textit{iri}.
}
traditional 3rd conjugation) had stem stress in the infinitive. As in Std. Italian, this stress distinction is maintained in Sicilian. Proposals involving a third Sicilian conjugational class consider verbs with antipenultimate stress to be a separate conjugation, so there are verbs in –ári, –íri (with penultimate stress) and –iri (with antipenultimate stress). All of these attempts fall a bit short, as there are definitely more than three ways of conjugating verbs in Sicilian (which all sources, of course, admit). For example, as in Balkan Romance (see Chapter 4, as well as the section on Aromanian later in this chapter), some Sicilian verbs in –iri have a stem extension from Latin [ESC] in certain forms. Plus, as in other Romance languages, some of the stem inconsistencies seen in present – perfect linkages are maintained, which again adds to the number of actually occurring conjugational patterns in Sicilian.

Again, these sources all talk about “the Sicilian dialect”, with little commentary on differences between the dialects of Sicily (Pitrè & Wentrup 1995 is an exception, but their discussion on variation in verb conjugation is quite brief). There are differences across the island of which speakers are aware⁴. Even within small areas, speakers who have had contact with speakers from other villages know that they speak differently; several of my informants commented that “we” conjugate a certain verb following a certain pattern, while “they” (people from a neighboring village) use a different pattern. For this reason, I have chosen to discuss the conjugational system of a specific dialect of Sicilian, that of the small town of Montedoro in the center of the island. The Montedorese data I collected consists of mainly present indicative and perfect forms, along with a few past participles. The present subjunctive is not used in this dialect (a typical characteristic of many Sicilian dialects), and nor is the synthetic future⁵ typical of Standard Italian (a periphrastic construction is used in Montedoro, as well as most of

---

⁴ For example, during a festival in Montedoro, there was a performance by a comedian whose routine consisted of several “Palermitani talk like this, and Catanesi talk like this”-type jokes. ⁵ Synchronically synthetic, diachronically analytic
Sicily/Southern Italy). This dialect has the imperfect indicative, but its specific patterns were not elicited as it appeared to be quite straightforward, as in Standard Italian and many other Romance languages.

6.3. **Verb Patterns in Montedorese Sicilian**

6.3.1. **Case 1: Present stem alternations**

When looking at infinitives, the trend described above is correct in Montedoro – there are infinitives in –ari and infinitives in –iri. However, these classes are not homogenous once their whole conjugational pattern is examined. As in Latin, it is not infrequent that verbs having the same type of infinitives often show striking differences, e.g., in linkages between present and past stems, etc. This can be even be seen in the present stems, where phonological complication has led to a situation where fragmentation has occurred.

Some verbs show what I will call a “typical” (or “non-alternating”) case of –ari verb conjugation. In this type of verb, there is no stem alternation, e.g., the verb cantari ‘sing’, has the stem cant(a)– throughout.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cântu</td>
<td>cantámmu</td>
</tr>
<tr>
<td>2</td>
<td>cânti</td>
<td>cantáti</td>
</tr>
<tr>
<td>3</td>
<td>cânta</td>
<td>cântanu</td>
</tr>
</tbody>
</table>

Table 6.1: PRES.IND paradigm of Sic. cantari ‘sing’

This is the pattern observed in the present indicative of a–stems when the stem vowel is a (e.g., cânta–). When there is a different vowel in the root, there is often a different

---

6 I could definitely be falling into the “this looks simple, so it has to be regular” trap, as seen in linguists’ reluctance to investigate Romance a–stem verbs in detail because they were (somewhat) regular in Latin, and look regular on the surface.

7 The orthography used here is based, more or less, on written Sicilian (on the rare, but increasingly common, instances, when it is written). Even though it is usually not written, word accent has been marked throughout.
pattern observed due to the existence of root allomorphy. For example, the following verb apparently has a stem vowel /o/:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>puássu</td>
<td>pussámmu</td>
</tr>
<tr>
<td>2</td>
<td>puássi</td>
<td>pussátu</td>
</tr>
<tr>
<td>3</td>
<td>póssa</td>
<td>póssanu</td>
</tr>
</tbody>
</table>

Table 6.2: PRES.IND paradigm of Sic. *pussári* ‘put’

The following verb, while also apparently having the stem vowel /o/, shows a different stem alternation pattern:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cóntu</td>
<td>cuntámmu</td>
</tr>
<tr>
<td>2</td>
<td>cónti</td>
<td>cuntáti</td>
</tr>
<tr>
<td>3</td>
<td>cúnta</td>
<td>cúntanu</td>
</tr>
</tbody>
</table>

Table 6.3: PRES.IND paradigm of Sic. *cuntári* ‘count’

As shown above, in the present indicative forms of *pussári* and *cuntári*, there is no difference in verbal endings, nor is there some sort of stem extension (e.g., [ESC]/[EDZ] discussed in Chapter 4, the velar augment discussed in Chapter 5). What does occur is a difference in vowel quality in the stem of the verb in different paradigm slots. In order to understand what is going on here, a basic background in (Montedorese) Sicilian phonology is needed. In stressed position, Sicilian (like Spanish, but unlike Standard Italian) has a typologically frequent five-vowel system. However, in unstressed position, there is only a three-vowel system, as unstressed *e* becomes *i* and unstressed *o* becomes *u*.

---

8 Unlike Spanish, the Sicilian mid vowels are consistently open [ɛ, ɔ]
9 Similar phenomena are seen frequently in other languages, both in Romance (e.g., Portuguese, Catalan, some Aromanian dialects) and outside of Romance (e.g., Greek dialects, Macedonian, etc.).
If we assume that the stem vowel is o in both of the verbs above (pussári and cuntári), then these verbs clearly follow different patterns. The 1PL and 2PL both have a stem vowel u as expected, as stress falls on the ending, rather than on the stem vowel in these forms (e.g., pussámmu, cuntámmu). However, the rest of the paradigms of these two verbs show differences. In pussári, the 1SG and 2SG stem vowel has been diphthongized to ua (e.g., 1SG puássu), while the stem vowel in the 3SG and 3PL is o (e.g., 3SG pössa). This diphthongization in the 1SG and 2SG is a case of metaphony, essentially a type of vowel harmony, that occurs when the following vowel (even if there are intervening consonants) is high.

In cuntári, the 1SG and 2SG show o (which would be expected given a stem vowel o, e.g. 1SG cóntu). However, the stem vowel is raised in the 3SG and the 3PL. The raising of this vowel is unexpected given that this vowel is stressed in these forms (3SG čúnta, 3PL cúnganu). This is not an isolated pattern, as it occurs in other verbs as well.

The question that must be asked is if this is predictable or not. In some ways, it is, as similar phonological phenomena can be seen elsewhere in the language. However, speakers still have to know which verbs are associated with each pattern. Definitely, one way to account for the different patterns would be to propose that the differences between the patterns of pussári and cuntári is due to their having different underlying stem vowels. However, what would the identity of these underlying vowels be? It could be proposed that the underlying stem vowel of pussári is o, which is raised in unstressed position and diphthongizes to ua before a high vowel. Then, what is the underlying stem vowel of cuntári? It is likely not u as there are verbs with stem u that maintain it throughout the paradigm:

---

10 These patterns can be seen elsewhere in the language, for example, in plural formation, e.g., piru 'pear-SG', pera 'pear-PL'; uassu 'bone-SG', ossa 'bone-PL' or in gender marking (e.g., biaddu 'pretty-MASC.SG', bedda 'pretty-FEM.SG'.

179
Thus, the underlying theme vowel of *cuntāri* would have to be something other than /u/ or /o/. However, there really are no other suitable candidates. It looks like that there are just two different alternation patterns and speakers have to memorize which verbs follow each pattern.

The same situation can be seen in verbs with a front stem vowel. Some show an alternation pattern involving diphthongization in the 1SG and 2SG (similar to *pusséri*), while some show the 1SG and 2SG having a high vowel and the 3SG/PL having a mid vowel (similar to *cuntāri*).

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>jiáttu</td>
<td>jittámmu</td>
</tr>
<tr>
<td>2</td>
<td>jiátti</td>
<td>jittáti</td>
</tr>
<tr>
<td>3</td>
<td>jétta</td>
<td>jéttaru</td>
</tr>
</tbody>
</table>

Table 6.5: PRES.IND paradigms of Sic. *jittāri* ‘toss’ and *tirāri* ‘throw’

Similar patterns are seen in verbs in *-iri*. Again, when the stem vowel is *a* there is no stem alternation.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>páru</td>
<td>pariémmu</td>
</tr>
<tr>
<td>2</td>
<td>pári</td>
<td>pariti</td>
</tr>
<tr>
<td>3</td>
<td>pári</td>
<td>pärinu</td>
</tr>
</tbody>
</table>

Table 6.6: PRES.IND paradigm of Sic. *parīri* ‘seem’
In addition, some verbs with stems in front or back vowels do not show stem alternation, as can be seen in the following examples (we can assume these to be a u-stem verb and an i-stem verb, respectively).

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cúrru</td>
<td>curriému</td>
</tr>
<tr>
<td>2</td>
<td>cúrri</td>
<td>curríti</td>
</tr>
<tr>
<td>3</td>
<td>cúrrí</td>
<td>currínu</td>
</tr>
</tbody>
</table>

Table 6.7: PRES.IND paradigms of Sic. *cúrriri* 'run' and *mintíri* 'put'

There are other verbs that show the above alternations as well. The patterns are the same as in *jittári* and *pussári* above, but the diphthongized stem is also present in the 3PL form.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>liággíu</td>
<td>liggiémmu</td>
</tr>
<tr>
<td>2</td>
<td>liággi</td>
<td>liggíti</td>
</tr>
<tr>
<td>3</td>
<td>léggi</td>
<td>liágginu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>duármu</td>
<td>durmiémmu</td>
</tr>
<tr>
<td>2</td>
<td>duármi</td>
<td>durmiti</td>
</tr>
<tr>
<td>3</td>
<td>dörmi</td>
<td>duárminu</td>
</tr>
</tbody>
</table>

Table 6.8: PRES.IND paradigms of Sic. *leggíri* 'read' and *durmiri* 'sleep'

This probably is due to the fact that in *-iri* verbs, the 3PL begins in a high vowel (*-inu*, instead of *-anu*, as in the verbs in *-ari*); we saw earlier that this diphthongization process occurred when the following vowel was high, as in the 1SG/2SG of verbs in *-ari*. The lack of stem diphthongization in the 1PL/2PL could be accounted for by saying that this does not occur in unstressed syllables (as stress falls on the ending, rather than the stem in the 1/2PL).

However, this does not account for the lack of diphthongization in the 3SG forms. The only difference between the 2SG and 3SG forms in the two verbs in Table 6.8, above, is that the verbal stem is diphthongized in the 2SG. Etymologically, the 3SG ending is *-e*, which is then reduced to *-i* as it is in unstressed position. Phonologically,
this could be accounted for by saying that the 3SG.PRES ending is underlyingly /e/.

There would be a synchronic rule that causes diphthongization before a high vowel, after which unstressed vowel raising occurs. However, since this 3SG ending for verbs in –iri always surfaces as –i, is there really any reason to believe that it is underlyingly anything else? It could allow us to explain why diphthongization does not occur in the 3SG, but it comes at the cost of proposing an underlying segment that never surfaces as such11 (cf. issues with absolute neutralization, Kiparsky 1968, 1971). Thus, just like the verbs in –ari, there is some unpredictability in the stem vowel pattern a verb will follow in –iri verbs (this is similar to the case of Spanish stem-changing verbs briefly discussed in Chapter 1).

In a traditional account, data such as those above would definitely warrant the proposal of the split of the traditional conjugational classes of Sicilian into subclasses based on the presence of these stem alternations. Even if we ignore traditionally-described “irregular verbs”, these examples show there is more to Sicilian verb conjugation than a simple split into an –ari class and an –iri class. Under the model proposed here, the data above warrant the establishment of a number of conjugational patterns. Again, as described in the previous chapters, these patterns are independent (i.e., one is not subsumed by another), though they are organized by the features they share. The data clearly demonstrate that the theme vowel is not the only aspect of conjugation that these speakers care about, as the presence of these vowel alternations also plays a role in verb categorization.

11 The only speakers that might have knowledge of this are those who know that the Italian 3SG ending in non–1st conjugation verbs is –e. Since Sicilian dialects are often thought as being a deviation from Italian, such speakers could potentially make this connection. However, it is not true that Sicilian is a deviation from Standard Italian, as these dialects formed simultaneously. In addition, generations of Sicilians have lived without any influence from Standard Italian, so there is no reason to think that the 3SG ending in these verbs is anything but –i.
6.3.2. Case 2: Verbs in \(-diri/ -dari\)

Another interesting aspect in the verbal morphology of this Sicilian dialect involves verbs that have the consonant \(d\) before the theme vowel (e.g., \(gridári\) ‘yell-INF’, \(crídiri\) ‘believe-INF’). When conjugated, the \(d\) is absent in the 1SG, while present throughout the rest of the present paradigm.

<table>
<thead>
<tr>
<th>INF</th>
<th>1SG.PRES.IND</th>
<th>2SG.PRES.IND.</th>
</tr>
</thead>
<tbody>
<tr>
<td>crídiri</td>
<td>créju</td>
<td>crédi</td>
</tr>
<tr>
<td>gridári</td>
<td>gréju</td>
<td>grédi</td>
</tr>
<tr>
<td>vidíri</td>
<td>víju</td>
<td>vidi</td>
</tr>
</tbody>
</table>

Table 6.9: Montedorese verbs in \(-diri/ -dari\)

Given the above data, one might think that there could be a synchronic rule that weakens \(d\) to \(j\) when followed by \(u\) (however phonetically odd this rule seems).

Potentially additional evidence is the 1SG of ‘go’ \(váju\). The verb is quite irregular and suppletive, but when compared to Standard Italian 1SG \(vádo\), it looks like the same process may be at work. However, one speaker gave an alternate form \(vídu\) for the 1SG of ‘see’. This could be seen as potentially influence from Italian (cf. Std. It. \(védo\) ‘see–1SG’), or perhaps due to paradigm leveling. In either case, it shows that this potential \(d\)-weakening process is not completely phonological, at least for this specific speaker. Even if this were influence from Italian, we cannot just set it aside as an “Italianism”.

This verb form is part of the speaker’s verbal system, regardless of its origin.

More interestingly, all the interviewed speakers gave the following forms for the verb \(rido\) ‘laugh’:

<table>
<thead>
<tr>
<th>INF</th>
<th>1SG.PRES.IND.</th>
<th>2SG.PRES.IND.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ridíri</td>
<td>rëdu</td>
<td>rëdi</td>
</tr>
</tbody>
</table>

Table 6.10: Montedorese \(rido\) ‘laugh’
No speaker interviewed gave a 1SG form réju, which one might expect given the trends seen in créju, váju, etc. This demonstrates that this conjugational information must be lexically-specific. This is not a regular phonological process, and verbs like ridiri that show d in the 1SG form must be categorized differently from verbs like cridiri in which it is absent.

6.3.3. Case 3: Verbs with \[ESC\]12

Just as in Balkan Romance and several other Romance languages, certain Sicilian verbs in -iri show a reflex of the Latin suffix [ESC] in certain paradigmatic slots, e.g., in the verb capíre 'understand':

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>capísci</td>
<td>capiémmu</td>
</tr>
<tr>
<td>2</td>
<td>capísci</td>
<td>capiti</td>
</tr>
<tr>
<td>3</td>
<td>capísci</td>
<td>capiscinu</td>
</tr>
</tbody>
</table>

Table 6.11: PRES.IND paradigm of Sic. capiri 'understand'

Similar to Catalan, Daco-Romanian and Std. Italian, the stem extension occurs in the singular and 3PL of the present the paradigm. Clearly, not all verbs in -iri show this stem extension (several of the examples discussed in Section 6.3.1. and 6.3.2. are -iri verbs that do not show it). I have not found any examples of any additional paradigmatic extension (e.g., an extension to the 1PL, or to other tenses, etc.) in this dialect.

There are some cases that show variation in that there are some verbs where one speaker shows a form with the stem extension, but another speaker shows the same verb without the stem extension. This is seen in the verb patiri 'suffer':

12 As shown in the examples, the vowel in this suffix is i in this variety of Sicilian (similar to Standard Italian and French) instead of e (as in Catalan and Balkan Romance). For convenience, I will continue to call this suffix [ESC].
While Speaker 1 does not show the stem extension in this verb, Speaker 2 does. One speaker that lacked the stem extension in this verb told me that younger people have [ESC] in lots of verbs that older speakers do not, due to influence from Standard Italian (e.g., Std. It. patisco ‘suffer–1SG.PRES.IND’). Situations like that above are not rare in Montedoro. My informant’s explanation for this is probably correct, though it need not concern us here. For us, all that matters is that these two speakers categorized this verb differently.

6.3.4. Case 4: Other present stem “irregularities”

In addition to the issues discussed above, there are several verbs that deviate even further from the basic patterns. Some of these are wildly irregular, showing suppletion within the present system.

<table>
<thead>
<tr>
<th>jíri ‘go’</th>
<th>éssiri ‘be’</th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>1  vaju</td>
<td>jammu</td>
</tr>
<tr>
<td>2  va(i)</td>
<td>jiti</td>
</tr>
<tr>
<td>3  va</td>
<td>vannu</td>
</tr>
<tr>
<td>1  sugnu</td>
<td>siemu</td>
</tr>
<tr>
<td>2  sei</td>
<td>seti</td>
</tr>
<tr>
<td>3  e</td>
<td>sonnu</td>
</tr>
</tbody>
</table>

Table 6.13: PRES.IND paradigms of Sic. jíri ‘go’ and éssiri ‘be’

There are other verbs that show “irregularities”, though they are too numerous to discuss fully here. While these verbs are normally referred to as being “irregular”, some do share characteristics with one another. For example, there is a group of verbs that,
like sugnu 'be–1SG', above, show a 1SG in –gnu, e.g., dugnu 'give–1SG.PRES.'; tiagnu 'hold–1SG.PRES.', viagnu 'come–1SG.PRES'. The rest of the conjugation of these verbs definitely varies, which shows that they pertain to, according to the principles laid out earlier, different conjugational patterns. However, the features that they do have in common may prove important in their categorization.

Another tendency that can be observed is that verbs that have a 1SG in –gnu or have other “irregularities” in the 1SG, tend to have a 3PL in –nnu:

<table>
<thead>
<tr>
<th>1SG.PRES.IND</th>
<th>3PL.PRES.IND</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>súgnu ~ sónnu</td>
<td>‘be’</td>
<td></td>
</tr>
<tr>
<td>stáju ~ stánnu</td>
<td>‘be’</td>
<td></td>
</tr>
<tr>
<td>váju ~ vánnu</td>
<td>‘be’</td>
<td></td>
</tr>
<tr>
<td>viágnu ~ viánnu</td>
<td>‘come’</td>
<td></td>
</tr>
<tr>
<td>puázzu ~ puánnu</td>
<td>‘can’</td>
<td></td>
</tr>
<tr>
<td>sácciu ~ sánnu</td>
<td>‘know’</td>
<td></td>
</tr>
<tr>
<td>fázzu ~ fánnu</td>
<td>‘make’</td>
<td></td>
</tr>
<tr>
<td>vuáglu ~ vuánnu</td>
<td>‘want’</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.14: 1SG/3PL.PRES.IND of “irregular” Sicilian verbs

However, this trend does not always hold. For example, the 1SG.PRES of dári ‘give’ is dúgnu. Given the trends seen above, it would seem that the 3PL would be dánnu. While this was the form that one speaker gave, another speaker have dúnanu, which follows the more “regular” pattern of 3PL formation (cf. cántu ‘sing–1SG’, cántanu ‘sing 3PL’). In addition, the 1SG.PRES of téniri ‘hold’ is tiágnu. Given the above forms (especially viágnu, which looks very similar in most forms), the expectation would be something like tiánnu for the 3PL.PRES. form. The consensus across all speakers interviewed was that the 3PL was tiánnu, again following the more “regular” pattern (except we would probably actually expect tiáninu).

13 These have to do with the palatalization of the final stem consonant by a jod (from the Latin verbs that had a 1SG in –iō). Its distribution is a bit unexpected, but since speakers have no idea why this was caused, they are not considered to be phonologically predictable.
Just because these verbs are “irregular” does not mean that they do not share some characteristics with each other, as well as with the other verbs discussed earlier. They are different patterns, and must be categorized as such. However, these similarities may be telling in how speakers actually organize verbs.

6.3.5. Case 5: Perfect system patterns and present–perfect linkages

Generally speaking, there are two types of perfective past tense in Romance, one synthetic (e.g., Std. It. cantai ‘I sang’, known in Italian as the passato remoto) and one compound (e.g., Std. It. ho cantato ‘I sang/I have sung’, known in Italian as the passato prossimo). While Standard Italian (and most dialects of north/central Italy) favor the passato prossimo, with the passato remoto rarely used in spoken conversation, Southern Italian dialects (including Sicilian) traditionally use the passato remoto exclusively. Due to recent Standard Italian influence, the passato prossimo is occasionally used in Sicilian, but it was not given attention in my fieldwork.

In Montedoro Sicilian, the perfect tense follows two general patterns – a majority pattern (which I call the single-stem pattern) that is followed by nearly all verbs in –ari and most verbs in –iri; and a pattern with a more limited lexical range (actually, a variety of patterns that share certain characteristics) that occurs most often in certain –iri verbs (this is referred to as the dual-stem pattern). The verbs that show the dual-stem pattern are reflexes of Latin verbs that showed different stems in the present and perfect systems. While this explains which verbs have this pattern diachronically, it does not explain how speakers categorize these in Modern Sicilian.

The following examples illustrate the single-stem pattern of the perfect tense in Montedorese Sicilian:

---

14 There is a semantic difference, but since most dialects favor one over the other in spoken communication, each can fulfill both roles.
Verbs conjugated in this manner are rather consistent. The only inconsistency is in the 3PL, where it appears that some –iri verbs have the ending –ieru, while others have –iru, e.g., the 3PL.PERF.IND. of ‘open’ appears to be grapíru, rather than grapiéru, and the 3SG.PERF.IND. of ‘read’ seems to be liggíru, rather than liggíeru. The actual distribution needs to be investigated in more detail to see if this is a phonological process at work.

An analysis of the perfect system becomes more complicated when observing the verbs that follow the dual–stem pattern. The general pattern in these perfects is that there are two different stems; one in the 1\textsuperscript{st} and 3\textsuperscript{rd} person forms and the other (which is the same as the present stem) in the 2\textsuperscript{nd} person forms.

For example, the perfect paradigm of the verb muviri ‘stay’ shows two stems, the present stem muvi– in the 2SG/PL, and the stem mossi– (which alternates with muassi–) in the rest of the paradigm.

---

15 Those familiar with Standard Italian might note that this looks like the Italian 1SG imperfect ending –avo, e.g., Std. It. cantavo ‘sing–1SG.IMP.IND.ACT.’. The Italian 1SG imperfect is etymologically –ava, and final –o has been extended from other 1SG forms (e.g., present, future) that have final –o. Montedorese (but not all Sicilian dialects) has maintained 1SG imperfect –ava, but has brought in –avu in as the 1SG perfect ending. The Montedorese imperfect endings have undergone an interesting evolution as well; see Costanzo (forthcoming) for details.
This distribution (which is similar, but not the same as found in the Standard Italian passato remoto\textsuperscript{16}) is fairly consistent, though I have come across a few examples that have a slightly different distribution, as discussed below.

Again, as in Latin and in the rest of Romance, the verbs in –ari are rather regular in the perfect, and few verbs form this alternative perfect (see discussion on Latin perfects in Chapter 2). There are a few exceptions, such as those below.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{dári ‘give’} & \textbf{stári ‘be’} \\
\hline
\textbf{singular} & \textbf{plural} & \textbf{singular} & \textbf{plural} \\
\hline
1 & détti & diáttimu & 1 & stétti & stittimu \\
2 & dásti & dástivu & 2 & stásti & stastivu \\
3 & détti & diáttiru & 3 & stétti & stittiru \\
\hline
\end{tabular}
\caption{PERF.IND paradigms of Sic. dāri ‘give’ and stāri ‘be’}
\end{table}

For ‘be’ one speaker also gave the 1SG.PERF. form stivu. (following the dominant pattern, but as if it were an –iri verb). The rest of the paradigm was as shown above. Perhaps the motivation behind this is to distinguish the 1SG and 3SG perfect forms, which are normally identical in these dual-stem perfects. However, the same speaker did not give a form dívu or dávu for the 1SG of ‘give’.

There are many –iri verbs that show a dual-stem perfect. Most show the same pattern (that shown in Table 6.16) in terms of the distribution of the two stems. However, again, there are some exceptions. For example, one respondent conjugated the perfect of scríviri with the two different stems. However, the 1SG and 1PL follow the

\textsuperscript{16} In addition to having STEM B in the 2SG/2PL, Standard Italian also has STEM B in the 1PL.
dominant perfect pattern (showing the present stem), with only the 3SG and 3PL showing the alternative stem.

<table>
<thead>
<tr>
<th>SPEAKER 1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
<td></td>
</tr>
<tr>
<td>1 scrivivu</td>
<td>scriviémmu</td>
<td></td>
</tr>
<tr>
<td>2 scrivisti</td>
<td>scrivistivu</td>
<td></td>
</tr>
<tr>
<td>3 scríssi</td>
<td>scríssiru</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.19: PERF.IND paradigm of Sic. scriviri 'write', speaker 1

When asked if the 1SG could also be scríssi, the informant said that it could, though they found it to be too “formal”. Perhaps they were associating this form with Standard Italian. Another speaker conjugated this verb completely according to the dominant pattern:

<table>
<thead>
<tr>
<th>SPEAKER 2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
<td></td>
</tr>
<tr>
<td>1 scrivivu</td>
<td>scriviémmu</td>
<td></td>
</tr>
<tr>
<td>2 scrivisti</td>
<td>scrivistivu</td>
<td></td>
</tr>
<tr>
<td>3 scriví</td>
<td>scrivíru</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.20: PERF.IND paradigm of Sic. scriviri 'write', speaker 2

There are other similar cases of variation, where one speaker conjugated a perfect paradigm under the dual–stem pattern (to some degree) while another conjugated it according to the single–stem pattern. The following two examples come from the same two speakers who gave the variant forms of scríviri, above. Speaker 1 (to an extent) used the dual stem pattern and Speaker 2 used the single–stem pattern for the perfect of scríviri. However, the opposite is true for putiri ‘be able’.
Here, Speaker A conjugated this verb under the single-stem pattern. However, he did note that pótti (the expected 3SG under the dual-stem pattern) is also acceptable for the 3SG. Speaker B conjugated this under the dual-stem pattern. The one difference is that the 1PL shows the present stem (we would expect puáttimu, if this was according to the dual-stem pattern completely).

Even if all the dual-stem patterns were completely consistent in the distribution of the two stems, it would still not be possible to divide perfect tense formation patterns into two groups just based on whether the perfect follows the single-stem of the dual-stem pattern. The relationship between the two stems in the dual stem pattern is not predictable. There are different correspondences between the two stems in different verbs.
Some of these look fairly similar, e.g., *detti, mitti, stetti*, etc. However, these perfect tense patterns are still not predictable from the present tense. The fact that they show dual-stem perfects is the only feature they share. These cannot even be categorized together because, even though they share the feature of having a dual-stem perfect, the actual details of the formation of the perfect (i.e., the linkages between present and perfect) are distinct.

### 6.3.6. Sicilian – conclusions

The data presented in the preceding sections show five cases in which a simple conjugational system of two (or three or four) verb classes does not account for all the data. If Sicilian speakers really cared about, over anything else, the vowel of the infinitive when conjugating verbs, the above examples would not exist. Saying that there are two main classes and all of the patterns branch off of these implies that the theme vowel is the main factor in categorization recognized by speakers. These data clearly show that there is more at play here then just the theme vowel of the verb in question when it comes to categorization.
6.4. BACKGROUND ON AROMANIAN AND AROMANIAN CONJUGATION

6.4.1. Aromanian in the Republic of Macedonia

Aromanian is a Balkan Romance language spoken in pockets throughout the Balkan Peninsula, principally in Greece and Albania, but also by smaller, but significant, populations in Macedonia, Serbia, Bulgaria, and Romania. The actual number of Aromanian speakers is difficult to estimate\(^\text{17}\), as it is spoken throughout the entire Balkan Peninsula, and census data are often incredibly accurate, particularly when dealing with linguistic minorities.

While the population of Aromanians in the Republic of Macedonia is estimated as being considerably lower than that of other Balkan countries, Aromanians in Macedonia are granted a far greater number of language rights than those in neighboring countries. For instance, some official paperwork is printed in Aromanian (along with the other minority languages), and there is some programming in Aromanian on the national television and radio stations. Aromanian also has co–official status in the municipality of Kruševo, where a substantial percentage of the population uses Aromanian as a native language. This diverges sharply from the situation in neighboring countries, where Aromanian (as well as other languages spoken by significant linguistic minorities) lacks such status.

An essential issue that must be addressed here is origins of Aromanians in the Republic of Macedonia. It appears that most of the Macedonian Aromanians originate from the southeast (near the present–day Albanian and Greek borders, or in present–day Albania or Greece). Most Aromanians in Macedonia claim to be descended from the Aromanians of Moscopole (the modern Albanian town of Voskopoja), a once thriving city that was nearly completely abandoned/destroyed when attacked repeatedly by Ottoman forces in the late 18th century (Winnifrith 1987). The population of Moscopole, which

\(^{17}\) The number of Aromanian speakers has a wide range, depending on which source one consults, some sources claim over 500,000, others claim under 100,000.
was once as high as 50,000, dispersed throughout the Balkans, many settling in what would become the modern Republic of Macedonia. From the various areas within Macedonia in which they settled, many Aromanians have moved to urban areas, e.g., Skopje, Bitola, Štip, etc., during the past century.

The municipality of Kruševo is usually considered to be the center of modern Aromanian culture in Macedonia. While there are fewer than one thousand speakers of Aromanian in Kruševo, they make up a decent-sized percentage of the population (the municipality has total population under 10,000). The dialect of Kruševo has been discussed in detail in Gołąb (1984), which deals with all aspects of the dialect, including some an overview of verbal morphology. There are also significant numbers of Aromanians living in Skopje and Štip, though an issue in these areas is that the Aromanian population (definitely in Skopje, and to a lesser extent, in Štip), has migrated to these locations from different areas, likely resulting in dialect mixing.

Bitola was (and to a lesser extent, still is) one of the major centers of Aromanian culture in the Balkans. While the population of Aromanians has greatly decreased over the past hundred years, the Aromanians are still a visible (as well as apparently rather powerful) part of Bitola society. There is an area of the city that was described to me as the “Aromanian neighborhood” directly west of the center, complete with an Aromanian church (although services appear to now be conducted in Macedonian). Although this neighborhood is gradually losing its Aromanian character, it is still home to some Aromanian speakers. Bitola is also home to one of the major Aromanian societies of Macedonia, the Comuna Armănească Frats Manakia, which is active in promoting the language and culture, and organizes events and publishes books, poetry, and a monthly newsletter, all in Aromanian.

---

18 Though there are no (as far as I am aware) studies on the Aromanian dialect(s) of Štip or Skopje.
One of the major issues with doing research on Aromanian in Bitola (or anywhere, for that matter), is that many of the Aromanian speakers living in Bitola originate from different Aromanian-speaking villages outside of the city. Even though all but one of the speakers I interviewed were born in Bitola, and many of their families have been in Bitola for several generations, the dialect features of their families’ villages of origin have been maintained to some extent. This is due to the fact that Aromanian is passed down within the family and not used to any great extent in public. Thus, the dialect(s) of Aromanian discussed here do not represent a so-called “Bitola dialect of Aromanian”, because, while all the informants were either born in Bitola or moved there at a young age, their families originate from different villages (e.g., Malovište, Gopeš, Trnovo, Nižepole) throughout the surrounding area.

In addition, it turns out to be difficult to even discuss a “Bitola dialect” because there are so many “layers” of Aromanian settlement of the city (similar to what has been described for the Ohrid–Struga region by Markovik’ 2003). Even though all the speakers I interviewed claim to descend from families originally from Moscopole, there was likely a substantial amount of dialectal variation within the Moscopole region. In addition, Aromanians of Moscopolean origin arrived in Bitola in different eras; some may have come directly from Moscopole to Bitola, some may have passed through several villages during the intermediate centuries, and so on.

6.4.2. Background on Aromanian dialects

While there has been significant research done on Aromanian, the extent to which it has been studied pales in comparison to that of the other Romance languages. Balkan Romance is often, in general, pushed to the side by many Romance linguists and when it is mentioned, the discussion is normally restricted to Standard Daco-Romanian (i.e., “Romanian”). Still, there have been numerous excellent publications; early work includes
Weigand (1888) and Capidan (1932), which give basic overviews of the language, along with some information on verb morphology. Papahagi (1963) serves as both an Aromanian–Romanian/French dictionary as well as an Aromanian etymological dictionary. Vrabie (2008) is an English–Aromanian dictionary that includes a wealth of information on variation within Aromanian, and often indicates differences (mainly lexical, but morphological information can be gained from this) between the main varieties of Aromanian. In addition, there are Aromanian–Macedonian dictionaries published by the major Aromanian organizations of Macedonia (Cuvata 2006, 2009, Popnicola 1997). Dealing with a specific dialect of Aromanian spoken in Macedonia, Golab (1984) is a description of the dialect of Kruševo, discussed above, while Markovik (2007) provides a description of the dialect spoken in the Ohrid–Struga region (east of Bitola, near the Macedonian–Albanian border).

6.4.3. Traditional views of Aromanian conjugation

As is the traditional view of Daco–Romanian, most sources (Weigand 1888, Wace & Thompson 1913, Capidan 1932, Vrabie 2000, Ianachiesci–Vlahu 2001), claim that Aromanian has four conjugational classes, one (more or less) descended from each of the four traditional Latin conjugations. However, the precise features used to define and delineate the four proposed classes are different in the various descriptions. Among the earlier descriptions is that in Wace & Thompson (1913); while their work mainly focuses on the history and culture of Aromanians (most attention is given to populations in Northern Greece, but the whole Aromanian–speaking area is covered to a more limited extent), they provide a chapter on the Aromanian language. They, as do most traditional descriptions, describe four conjugational classes. Their description is based both on the infinitive and how the perfect tense is formed (which, when compared to any traditional account of any Romance language, is somewhat unique):
The first conjugation contains most verbs of the Latin first conjugation, verbs in *edzu* and *zburăscu* to speak.
The second conjugation makes the infinitive in *eare* and the preterite in *uî*; it contains verbs of the second and third Latin conjugations, a few of the fourth, and the few verbs in *−sku* which are of Latin origin like *pasku pășteare*. 19
The third conjugation mainly contains verbs of the Latin third conjugation and makes the infinitive in *eare* and the preterite in *u*. The fourth conjugation contains verbs of the fourth and third Latin conjugations, the infinitive in *ire* and the preterite in *ii*. To this conjugation can also be reckoned all the verbs in *esku* which are almost without exception foreign verbs incorporated in Vlach 20...

(Wace & Thompson 1913:243–246)

This description is different, not in terms of the number of conjugations, but in the principles of classification. Most of the traditional classifications base the separation into four classes solely based on the quality of the stem vowel and do not make any reference to the perfect. This is seen in, e.g., Ianachiesci–Vlahu (2001), which describes four distinct stem vowels in the present stem (rather than the three in Wace & Thompson). These vocalic differences can be clearly seen in the 2PL.PRES.IND.:

The specific suffixes are seen in in the 2nd person plural of the present indicative:

- *−a* (accented) in verbs of the 1st conjugation;
- *−e* (accented) in verbs of the 2nd conjugation;
- *−i* (unaccented) in verbs of the 3rd conjugation;
- *−i* (accented) and *−ă* (accented) in verbs of the 4th conjugation

(Ianachiesci–Vlahu 2001:97) 21

19 These verbs in *−sku* have the suffix extended throughout the paradigm. This is unlike the verbs of the traditional 4th conjugation (as in Daco–Romanian) which have a restricted paradigmatic appearance.
20 The term “Vlach” has a complicated history. It is an exonym; it is used by others of the Balkans to describe Aromanians (e.g., Grk. βλάχοι, Trk, ulahlar, Mk. vlasi, etc. ‘Aromanians’) but not used by Aromanians to describe themselves (they use armânji ‘Aromanians’). In English, much early work (e.g., Wace & Thompson 1913) used ‘Vlach’, but ‘Aromanian’ is much more commonly used today.
21 “Sufixurile spetsifite s–ved la a II–a persoană plural, indicative presuntu: *−a* (actsentuat) ti verbile di conjugarea I; *−e* (actsentuat) ti verbile di conjugarea a II–a; *−i* (niactsentuat) ti verbile di a III–a conjugare; *−i* (actsentuat) shi *−ă* (actsentuat) ti verbile di a IV–a conjugare.”
Like Ianachiesci-Vlahu, Vrabie (2000) proposes four Aromanain conjugations, but says that the differences can best be seen in the verbal noun (essentially the outcome of the Latin infinitive, but it is never used as an infinitive\(^{22}\) in Aromanian, e.g., Lat. \texttt{DORMIRE} ‘sleep-INF’ → AR \texttt{durňire} ‘sleeping-NOUN’.

In his description of the Aromanian dialect of Kruševo, Gołąb (1984) maintains the 4–conjugation system, but says that there really is no reason of keeping the conjugation II separate from (as he puts it, the “so-called”) conjugation III\(^{23}\):

I preserve the classification of traditional Rumanian grammars, but from a structural standpoint there is no distinction between the II and the so-called III conjugation: both have the same underlying (basic) present–tense stem in \{-e\}; the only difference is that the so-called III conjugation shows in plural the antepenultima stress, which entails an automatic reduction of the unstressed /e/ into /i/.

(\textit{Gołąb 1984: 114})

The essential difference between the descriptions of Ianachiesci-Vlahu, Vrabie, and Gołąb on one hand, and Wace & Thompson, on the other, is that Wace and Thompson do not consider there to be verbs that have a present stem in unstressed \textit{i}, and thus, in their main description, do not consider Aromanian verbs that have antepenultimate stress in the 1PL/2PL. Later on, Wace & Thompson do discuss the stress difference, but consider it is a dialectal feature that is found in “central” dialects and is due to Greek

\(^{22}\) Aromanian, like several other languages of the Balkans, lacks an infinitive. In its place, a particle \textit{s}– plus a conjugated form of the present subjunctive of the verb is used, e.g., AR \textit{pot s-beau} (can–1SG drink–1SG) ‘I can drink…’ (cf. Cat. \textit{puc beure} (can–1SG drink–INF))

\(^{23}\) A similar proposal for Daco–Romanian can be seen in Ruhlen (1974), which is subsequently critiqued in Juillard 1978.
influence (though later descriptions of northern dialects, e.g., Kruševo, do show this feature).

In the Ohrid–Struga dialect of Aromanian, Markovik’ (2007) claims that there are only three conjugations:

In the Aromanian dialect of Ohrid, verbs can be divided in three groups according to the vowel that ends the present stem. The full form in the stem best is seen in the second person plural. So, in the Aromanian dialect of Ohrid, the base can end in –a, –e, –i. We note that in most work on Aromanian, four verbs groups are listed. In the Aromanian dialect of Ohrid, the so-called third conjugation (according to traditional grammars) has passed to the second conjugation.

(Markovik’ 2007:109)

However, this is a particularity of this dialect, where, according to Markovik’, all the traditional third conjugation verbs (the conjugation showing stress differences in the 1/2PL) have moved to the second conjugation.

There is quite a bit of variation in Aromanian dialects in Macedonia as to the identities and inventories of these “conjugations”. Ianachiesci–Vlahu (2001), a grammar for students learning Aromanian, gives a series of paradigms listing the conjugated forms of irregular verbs and certain cases of variation as an appendix. Many of these verbs given vacillate between conjugations, particularly between the 2nd and 3rd conjugations. The only difference between the 2nd and 3rd conjugations (in the present tense) is in the 1PL and 2PL, where the endings are –ém/–éts, respectively, in the 2nd conjugation, and unstressed –im/–its, respectively, in the 3rd conjugation.

24 "Во ароманскиот охридски говор глаголите може да се поделат во три групи според вокалот на кој завршува презентската основа. Полната форма на основата најдобро се гледа во второ луѓе множина. Така во охридскиот аромански говор, основата може да завршува на –а, –е, –i. Мора да напоменеме дека во повеќето трудови за ароманскиот јазик се наведени четири глаголски групи. Во ароманскиот охридски говор таканаречената трета конјугација (според традиционалните граматики) сосема преминала во втората конјугација.”
This type of variation appears to be fairly common in some Aromanian dialects. In my research, I often saw variation between these forms.

Aromanian dialects show the same stem extensions that Daco–Romanian does: a [EDZ] in certain a-stem verbs and [ESC] in certain i-stem verbs. No Aromanian grammar consulted considers these two be separate conjugations. Rather, as in most Daco–Romanian descriptions (as well as descriptions of Catalan, Italian, etc.) they are just referred to different subclasses of the 1st (a-stem) and 4th (i-stem) conjugations. In addition, there is quite a bit of irregularity in the simple perfect, much of which is held over from the messiness of the Latin perfect.

The following section discusses several aspects of verb conjugation that I found in my fieldwork on Aromanian in Bitola. Due to the nature of the data, no conclusions can be made. However, the data, as fragmentary as it is, does show some interesting aspects of Aromanian verbal morphology.

6.5. Aspects of Verb Conjugation in “Bitola Aromanian”

6.5.1. Case 1: Present stem alternation, a-stems

The defining characteristics of verbs that have a present in –a (the traditional 1st conjugation discussed in the various descriptions above) include the fact that the 3SG

<table>
<thead>
<tr>
<th>3&lt;sup&gt;rd&lt;/sup&gt; conj. (stem ı, unstressed)</th>
<th>1SG</th>
<th>2SG</th>
<th>3SG</th>
<th>1PL</th>
<th>2PL</th>
<th>3PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>duc</td>
<td>duts</td>
<td>dutse</td>
<td>dûtsim</td>
<td>dûtsits</td>
<td>duc</td>
<td></td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; conj. (stem é, stressed)</td>
<td>duc</td>
<td>duts</td>
<td>dutse</td>
<td>dûtsém</td>
<td>dutsets</td>
<td>duc</td>
</tr>
</tbody>
</table>

Table 6.25: PRES.IND paradigm(s) of AR duc 'take' (Ianachiesci–Vlahu 2001)
and the 3PL are typically the same, both having the ending -ã. The 1PL and 2PL are also fairly straightforward, with endings -ãm and -ats, respectively. The 1SG and 2SG PRES.IND forms are a bit more complicated. When there is only one stem-final consonant, the 1SG has no ending and the 2SG is marked via palatalization of the final consonant. The example below shows this pattern:

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 mãc</td>
<td>mãcãm</td>
</tr>
<tr>
<td>2 mâts</td>
<td>mãcats</td>
</tr>
<tr>
<td>3 mãcã</td>
<td>mãcã</td>
</tr>
</tbody>
</table>

Table 6.26: PRES.IND paradigm of AR mãc 'eat'

When the stem-final consonant is already palatalized, or is r, the 2SG is not "re-palatalized", and the 1SG and 2SG are identical. For example, the 1SG of 'study' is nvets. Since the final consonant is not able to be palatalized, the 2SG is also nvets. The same is seen in verbs with final -r, e.g., 1SG adar, 2SG adar 'fix'.

If the stem ends in two (or more) consonants, the 1SG ending is -u and the 2SG ending is (normally) -i. The rest of the paradigm is as described above.

---

26 The same pattern can be seen in Daco–Romanian. I have found some cases of AR a-stems whose 3SG and 3PL are different. These cases will be discussed below.

27 Most consonants have a palatalized counterpart that occurs (usually) before i, sometimes before e and in certain grammatical forms. The following chart shows Aromanian consonants and their palatalized counterparts.

<table>
<thead>
<tr>
<th>non.pal.</th>
<th>pal.</th>
<th>non.pal.</th>
<th>pal.</th>
<th>non.pal.</th>
<th>pal.</th>
<th>non.pal.</th>
<th>pal.</th>
</tr>
</thead>
<tbody>
<tr>
<td>t, c</td>
<td>ts</td>
<td>m, n</td>
<td>nj</td>
<td>p</td>
<td>ch</td>
<td>v</td>
<td>y</td>
</tr>
<tr>
<td>d, g</td>
<td>dz</td>
<td>l</td>
<td>lj</td>
<td>b</td>
<td>gh</td>
<td>s</td>
<td>sh</td>
</tr>
</tbody>
</table>

Other consonants, such as r and consonants that are already palatalized (e.g., one of the palatalized variants above) do not have a palatalized variant.

28 Since Aromanian technically does not have an infinitive, the common citation form for Aromanian verbs is the 1SG.PRES.IND.
However, if the stem ends in two (or more consonants) and the final consonant of the stem is able to be palatalized in the 2SG (the example above is not able to be palatalized, as it has final -ɾ), then the final consonant is palatalized, but the ending is now -ã.

### Table 6.27: PRES.IND paradigm of AR acumpru 'buy'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>acumpru</td>
<td>acumprăm</td>
</tr>
<tr>
<td>2</td>
<td>acumpri</td>
<td>acumprats</td>
</tr>
<tr>
<td>3</td>
<td>acumpră</td>
<td>acumpră</td>
</tr>
</tbody>
</table>

### Table 6.28: PRES.IND paradigm of AR cântu 'sing'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cântu</td>
<td>cântăm</td>
</tr>
<tr>
<td>2</td>
<td>cântsă</td>
<td>cântats</td>
</tr>
<tr>
<td>3</td>
<td>cântă</td>
<td>cântă</td>
</tr>
</tbody>
</table>

Something similar happens in a-stem verbs that have a stem final -tr. In the 2SG, the t is palatalized (even with intervening -r-), and the ending is again -ã.

### Table 6.29: PRES.IND paradigm of AR intru 'enter'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>intru</td>
<td>intrăm</td>
</tr>
<tr>
<td>2</td>
<td>intrsă</td>
<td>intrats</td>
</tr>
<tr>
<td>3</td>
<td>intră</td>
<td>intră</td>
</tr>
</tbody>
</table>

This is not necessarily true for all Aromanian verbs. I have data from a different speaker (whose family originates from a different village) who, for the verb 'bark' had 1SG latru, 2SG latri (not something like latsră). There was insufficient data to determine if this variation is variation between speakers or an inherent property of these verbs.

Even though some of these patterns look a little strange, they were quite consistent across speakers (somewhat remarkably, given the fact that the speakers were
originally (ancestrally) from different villages in the area around Bitola). Similar to some of the Sicilian data discussed above, there is some stem alternation regarding simple vowels vs. diphthongs in some verb paradigms. When the final stem vowel is e or o, the general pattern in a-stem verbs is that in the 1SG/2SG there is a simple vowel (e or o), in the 3SG/3PL there is a diphthong (ea or oa), and in the 1PL/2PL there is a raised/reduced vowel (e→i, o→u). The two verbs below (elicited from the same speaker) show this pattern:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>scol</td>
<td>scolãm</td>
</tr>
<tr>
<td>2</td>
<td>scolj</td>
<td>scolats</td>
</tr>
<tr>
<td>3</td>
<td>scoalã</td>
<td>scoalã</td>
</tr>
</tbody>
</table>

Table 6.30: PRES.IND paradigms of AR scol 'beg' and aplec 'get up'

For some verbs, I found some variation. Below are the present paradigms of 'ask' given by two different speakers.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ntreb</td>
<td>ntribãm</td>
</tr>
<tr>
<td>2</td>
<td>ntregh</td>
<td>ntribats</td>
</tr>
<tr>
<td>3</td>
<td>ntreabã</td>
<td>ntreabã</td>
</tr>
</tbody>
</table>

Table 6.31: PRES.IND paradigm of AR (ã)ntreb 'ask'; speakers 1,4

Speaker 1 shows diphthongization in the 3SG and 3PL, as expected. However, speaker 4 does in the corresponding forms. This is not widespread for speaker 4 – several of their other examples do show diphthongization. Variation is seen in speaker 1 as well; they showed diphthongization in the 3SG and 3PL of 'ask'. This is not the case for all

---

29 This is really very similar to the Sicilian examples except for the fact that the diphthong was in the 1/2SG in Sicilian and the simple vowel was in the 3SG/PL. In Aromanian, it is the complete opposite.
verbs that could be potentially diphthongized. Below are the present paradigms of
'carry' given by two speakers, speaker 1 and speaker 3:

<table>
<thead>
<tr>
<th></th>
<th>SPEAKER 1</th>
<th></th>
<th>SPEAKER 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>singular</td>
<td>plural</td>
<td>singular</td>
</tr>
<tr>
<td>1</td>
<td>portu</td>
<td>portãm</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>portsã</td>
<td>pertats</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>portã</td>
<td>portã</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 6.32: PRES.IND paradigm of AR *portu* 'carry'; speakers 1,3

In this case, speaker (1) does not have a diphthongized stem in the 3SG/PL. On the
other hand, speaker (3) has a diphthong in the 3SG but not in the 3PL. This looks odd,
as the 3SG and 3PL of *a*-stem verbs are typically identical. This shows, at least for
speaker (3), that diphthongization is not a completely phonological process as the 3SG
and 3PL present the same environment for diphthongization to occur. It looks like
speaker (1) categorized this verb as non-diphthongizing. In 'carry', speaker (1)
diphthongizes in the 3SG/3PL and in 'ask' they do not. These are not the only two
patterns that speaker (1) has regarding diphthongization in *a*-stem verbs:

<table>
<thead>
<tr>
<th></th>
<th>SPEAKER 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>singular</td>
</tr>
<tr>
<td>1</td>
<td>gioc</td>
</tr>
<tr>
<td>2</td>
<td>giots</td>
</tr>
<tr>
<td>3</td>
<td>gioacã</td>
</tr>
</tbody>
</table>

Table 6.33: PRES.IND paradigms of AR *gioc* 'dance' and *scol* 'lift'; speaker 1

In both of these examples, speaker (1) has a diphthongal stem in the 3SG and a non-
diphthongal stem in the 3PL (similar to Speaker (3)'s conjugation of 'carry' shown in
Table 6.32 above\(^{30}\)). Thus, apparently there are three patterns (no diphthongization,

\(^{30}\) There are various potential explanations for this. Perhaps since both the 3SG and 3PL were the
same, speakers replaced the diphthong with a simple vowel to distinguish the two forms (the
reintroduction of the simple vowel *o* could be analogical to the rest of the paradigm). This could
3SG and 3PL diphthongized, only 3SG diphthongized) for this single speaker for a-stem verbs that have a stem vowel e or o. This all makes the traditional picture of the “1st conjugation” more complicated, at least for this speaker.

In the above examples, where a speaker shows differences in diphthongization in the 3SG and 3PL, the 3SG was always the form that was diphthongized. However, another speaker (with family origins of a different village) showed the opposite distribution in another verb.

<table>
<thead>
<tr>
<th>SPEAKER 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>singular</td>
<td>plural</td>
</tr>
<tr>
<td>1</td>
<td>vom</td>
</tr>
<tr>
<td>2</td>
<td>vonj</td>
</tr>
<tr>
<td>3</td>
<td>vomã</td>
</tr>
</tbody>
</table>

Table 6.34: PRES.IND paradigm of AR vom ‘vomit’, speaker 4

Besides the unexpected presence of –em (rather than expected –ãm) in the 1PL, this example is interesting for the fact that the 3SG shows a simple vowel and the 3PL shows a diphthong.

### 6.5.2. Case 2: Present stems in –e, –i

There are also verbs that have stems in e and verbs that have stems in i.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ascundu</td>
<td>ascundem</td>
<td>1</td>
<td>dormu</td>
<td>durnjim</td>
</tr>
<tr>
<td>2</td>
<td>ascundzã</td>
<td>ascundets</td>
<td>2</td>
<td>dornji</td>
<td>durnjits</td>
</tr>
<tr>
<td>3</td>
<td>ascundi</td>
<td>ascundu</td>
<td>3</td>
<td>dormi</td>
<td>dormu</td>
</tr>
</tbody>
</table>

Table 6.35: PRES.IND paradigms of AR ascundu ‘hide’ and dormu ‘sleep’

also be analogical to the verbs in the other traditional conjugations. Verbs with stem e or i normally have the 1SG and 3PL with the same form (instead of the 3SG and 3PL, as in a-stem verbs). In these verbs, only the 3SG is diphthongized, so perhaps this pattern could have influenced these verbs. The reasons behind this, while interesting, really aren’t of importance here. The important thing is that speakers show these patterns.
In the present tense, the only differences in the endings are in the 1PL/2PL, which are 
\(-im/-its\) in \(i\)-stems and \(-em/-ets\) in \(e\)-stems. One characteristic of these types of verbs 
is that the 1SG and the 3PL are normally identical (rather than the 3SG and 3PL, as in \(a\)-
stems). There are a few \(e/i\) stems whose 1SG and 3PL are not identical, but these are 
usually considered irregular verbs (and have additional idiosyncrasies), such as those 
below:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dau</td>
<td>dadem</td>
</tr>
<tr>
<td>2</td>
<td>dai</td>
<td>dadets</td>
</tr>
<tr>
<td>3</td>
<td>da</td>
<td>da</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>voi</td>
<td>vem</td>
</tr>
<tr>
<td>2</td>
<td>vrei</td>
<td>vrets</td>
</tr>
<tr>
<td>3</td>
<td>va</td>
<td>vor</td>
</tr>
</tbody>
</table>

**Table 6.36: PRES.IND paradigms of AR dau ‘give’ and voi ‘want’**

### 6.5.3. Case 3: Stem extensions [ESC] and [EDZ]

As in other Romance languages, there is more than one type of verb that shows the 
suffix [ESC] in this variety of Aromanian. First is the group of verbs that were inherited 
from Latin with the suffix already attached. In these verbs, [ESC] occurs throughout the 
paradigm (all tenses). These verbs are usually \(e\)-stems.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>crescu</td>
<td>crishtem</td>
</tr>
<tr>
<td>2</td>
<td>creshci</td>
<td>crishtets</td>
</tr>
<tr>
<td>3</td>
<td>creshci</td>
<td>crescu</td>
</tr>
</tbody>
</table>

**Table 6.37: PRES.IND paradigm of AR crescu ‘grow’**

In verbs such as *crescu*, the presence of the stem extension is consistent throughout the 
paradigm, e.g., the 1SG.PERF. of this verb is *criscui*. As in several other Romance 
languages, [ESC] is also used as a stem extension in certain \(i\)-stem verbs. As in 
standard Daco–Romanian, Italian, Catalan, etc., within the present paradigm it occurs in 
the singular and the 3PL (and absent in the 1/2PL).
This is the most frequent pattern. However, one speaker appears to have this suffix in an odd distribution, as is seen in their conjugation of the verb 'sew', shown below:

Table 6.38: PRES.IND paradigm of AR *pistisepscu* 'believe'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pistisepscu</td>
<td>pistisepm</td>
</tr>
<tr>
<td>2</td>
<td>pistisepshci</td>
<td>pistiseps</td>
</tr>
<tr>
<td>3</td>
<td>pistisepshci</td>
<td>pistisepscu</td>
</tr>
</tbody>
</table>

The paradigm above is interesting for a few reasons. First, this is one of the rare cases in which a non *a*-stem verb does not have the same form in the 1SG and 3PL. Second, and more significant here, the 3PL form shows the suffix –*escu*. As explained above, the most frequent (both in terms of this AR dialect and in Romance in general) pattern is to have the stem extension in the singular and the 3PL. However, in this example, it only appears in the 3PL form. The same speaker also the following example, which shows the same pattern:\footnote{This could have potentially been motivated by a desire to distinguish the 1SG and 3PL, which, as discussed above, are usually identical in *e*-stems and *i*-stems. This is not consistent, as this speaker has some *e*-*i*-stems with identical 1SG/3PL in the present.}

Table 6.39: PRES.IND paradigm of AR *cos* 'sew'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cos</td>
<td>cusem</td>
</tr>
<tr>
<td>2</td>
<td>cosh</td>
<td>cusets</td>
</tr>
<tr>
<td>3</td>
<td>coasi</td>
<td>cusescu</td>
</tr>
</tbody>
</table>

\footnote{A borrowing from Greek πιστεύω, aorist πίστεψα. As in the other Balkan Romance languages, this group (*i*-stem/\(+\text{ESC}\)) is highly populated by borrowed verbs (in this dialect, most of the borrowings are from Greek).}
In addition to the group of verbs in -escu, there is also a small group of verbs in -ãscu. This is roughly equivalent to the distinction between Daco–Romanian verbs in -i and verbs in -i discussed in Chapter 4. The most notable example is the very common verb meaning 'speak'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>zbrâscu</td>
<td>zbrâm</td>
</tr>
<tr>
<td>2</td>
<td>zbrâshci</td>
<td>zbrats</td>
</tr>
<tr>
<td>3</td>
<td>zbrashci</td>
<td>zbrâscu</td>
</tr>
</tbody>
</table>

Table 6.41: PRES.IND paradigm of AR zbrâscu 'speak'

The suffix is shown, as in the verbs in -escu, in the singular and the 3PL. However, the suffix is different here; -ãscu instead of escu. The singular and 3PL still show non-a stem endings, but the 1PL and 2PL show a-stem endings. Most sources consider this verb and the few others that behave similarly to be i-stem verbs that show a slightly different pattern than normal i-stems that show the suffix. At one point it was thought that the difference was phonologically predictable, where i-stem [+ESC] verbs with with stem final consonant –r get äscu, while all other i-stem [+ESC] verbs get –escu. This works for the two examples here, and it may have been true at some point in the history of Balkan Romance, but in modern Aromanian there are verbs with stem–final consonant –r have a 1SG in –escu (Golab 1984). Wace & Thompson (1913) propose that this verb is an a-stem verb that shows some odd behavior (see quote in section 6.5.3). A different speaker conjugated this verb using a pattern that looked like a mixture of the two seen above in –escu and –ãscu.
Table 6.42: PRES.IND paradigm of AR *zbrescu* 'speak'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>zbrescu</em></td>
<td><em>zbram</em></td>
</tr>
<tr>
<td>2</td>
<td><em>zbreshci</em></td>
<td><em>zbrits</em></td>
</tr>
<tr>
<td>3</td>
<td><em>zbrashci</em></td>
<td><em>zbrescu</em></td>
</tr>
</tbody>
</table>

In this example, the 1SG, 2SG and 3PL show *-escu*, while the 3SG shows *-ascu*. The 1PL looks like an *a*-stem, while the 2PL looks like an *i*-stem.

The other suffix discussed in Chapter 4 is also present in this dialect of Aromanian. As in Daco–Romanian, [+EDZ] verbs are *a*-stems.

Table 6.43: PRES.IND active paradigm of AR *lucredz* 'work'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>lucredz</em></td>
<td><em>lucrâm</em></td>
</tr>
<tr>
<td>2</td>
<td><em>lucredz</em></td>
<td><em>lucrats</em></td>
</tr>
<tr>
<td>3</td>
<td><em>lucreadzã</em></td>
<td><em>lucreadzã</em></td>
</tr>
</tbody>
</table>

However, there is a bit of oddness in some verbs for some speakers:

Table 6.44: PRES.IND paradigm of AR *construedz* 'build'

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>construedz</em></td>
<td><em>construim</em></td>
</tr>
<tr>
<td>2</td>
<td><em>construedz</em></td>
<td><em>construits</em></td>
</tr>
<tr>
<td>3</td>
<td><em>construeadzã</em></td>
<td><em>construeadzã</em></td>
</tr>
</tbody>
</table>

This speaker has the expected forms except in the 1/2PL, which show *i*-stem endings.

The rest of the paradigm follows the *a*-stem [+EDZ] pattern.
6.5.4. Aromanian – conclusions

The discussion above dealt only with a few issues in the present tense conjugation of Aromanian verbs. There is much more that needs to be discussed. For example, the simple perfect tense of Aromanian, like that of Sicilian or Latin, is not completely predictable from the present tense. The basic situation is that the traditional 1st, 2nd, and 4th conjugations generally form the perfect following one pattern, while the traditional 3rd conjugation has a rather different pattern that is not predictable given the present tense. Many verbs of the traditional 3rd conjugation also have alternate forms showing both patterns. A similar situation holds for the past participle. These issues are reserved for discussion at a later date.

Even though the data presented here covers only a small portion of the Aromanian verbal system, there is clear evidence that any classification into a small number of conjugational classes, as in the traditional accounts, is problematic. There is evidence that other issues are definitely at play in the categorization of these besides solely the theme vowel (e.g., diphthongization, presence of stem extensions, etc.).

6.6. Summary and Conclusions

This chapter has discussed issues in verbal conjugation in the Sicilian dialect of Montedoro and the Aromanian dialects of Bitola. The examples shown demonstrate that a simple, traditional, classification of verbs into a few conjugational classes is not adequate to describe these systems. Again, the theme vowel, while being important, is not the only factor in verb categorization. There are other factors at play, e.g. presence/absence of a stem extension, presence/absence of diphthongization, linkages between present and perfect stems, etc. As the theme vowels do not seem to be the only important factors in verb categorization in these dialects, the traditional manner of
describing the conjugational classes of these varieties breaks down. A model that is able to allow for these "local" generalizations is needed to describe this system.
CHAPTER 7:
CONCLUDING REMARKS

7.0. INTRODUCTION

The preceding chapters demonstrate how traditional classifications of the Romance verbal system into conjugational classes fall short and how a different model, one based on the Gross Detail Method, and thus focuses on an explicit representation of each and every difference in related forms, proves advantageous. This reference here to “representations” suggests a cognitive dimension that is explored in this concluding chapter.

7.1. ON ROMANCE CONJUGATIONAL CLASSES

7.1.1. Weaknesses of traditional classifications

The vast majority of descriptions of Latin and Romance “conjugational classes” are dependent, either wholly or primarily, on the notion of the theme vowel (most easily seen in infinitival forms). In Latin, this is not a particularly successful means of classification, as there clearly are many more observed patterns than just four or five. As in Latin, this manner of describing Romance conjugational classes is not successful, as it is a holdover from Latin, where it was not successful to begin with. In addition, while trends of “simplification” occurred in the development from Latin to Romance, the various Romance verbal systems have developed certain characteristics that make them even more difficult to describe using a system of a very limited number of conjugations.
The practice of identifying conjugations by the vowel they present in the infinitive (where the number of theme vowels essentially equals the number of conjugations) is pervasive in Romance. For example, as discussed in Chapter 4, the Daco–Romanian verb system is traditionally considered to have four conjugational classes, one descended from each of the traditional Latin conjugations (Rauta 1974, Daniluc & Daniluc 2000, etc.). There are several proposals that add a fifth conjugational class in Daco–Romanian, one that comprises the small group of verbs that have infinitives in –î (Julliand & Edwards 1971, Mallinson 1986, etc.). While the guidelines of the Gross Detail Method explicitly stated in Chapter 1 clearly require that these verbs be considered “separate” in some way, it appears that the only reason why this is often posited (and accepted) by researchers is that this distinction is seen in the infinitive (i.e., infinitives in –î vs. infinitives in –i). At the same time, other “complications” (e.g., verbs showing a stem extension [ESC] or [EDZ]) are rarely considered to warrant the establishment of a separate conjugational class by most Romance scholars. The reasoning behind this has nothing to do with the facts of the language; and rather, it is due to the fact that there is no distinction made in the theme vowel of the infinitives (e.g., all a–stem verbs have the infinitival vowel a, regardless if they show the stem extension or not).

No account of Latin or Romance conjugation would argue that there are three/four/five conjugations that are homogeneous throughout, as there are obviously more than three/four/five actual patterns in which verbs are conjugated in any of these languages. These accounts start by positing just a few conjugational classes necessarily involve a substantial amount of subbranching of these classes (and thus, the positing of

1 Interestingly, however, it appears as more Balkan Romance sources actually do posit that the traditional subclasses (e.g., a–stem/[+EDZ], a–stem/[−EDZ]) are independent conjugational classes (e.g., Ciompec, et al. 1985, Lombard 1955, Lombard & Gâdei 1981, etc.) than sources describing other Romance varieties. For example, I have yet to see a description of Italian, Catalan, or French that gives full conjugational class status to their respective traditional i–stem[+ESC] and i–stem[−ESC] subclasses. Despite the fact that this is more common in descriptions of Balkan Romance, it is still not the norm (see Ciompec, et al. 1981).
“subclasses”) in order to account for all the patterns attested in the language. Traditionally, the number of conjugational classes is determined by the number of different theme vowels verbs in that language have. However, is there any reason why the theme vowel should be thought of as the primary classificatory feature, i.e., the feature upon which the first instance of branching is based? No, it is just a well-established tradition that, despite its long establishment, does not have solid reasoning behind it. Even accounts that attempt to somehow systematically account for all the differences in inflection (e.g., Dressler 2003, Aguirre & Dressler 2006) are still somewhat biased by the theme vowel, in that in their first main split into “macroclasses” is still defined by the theme vowel (see Chapter 2).

The whole point of positing subclasses is that doing so captures a generalization regarding some feature that subclasses of a certain class share to the exclusion of other classes. This type of model is rather one-dimensional, and it could see potential success in classifying some group of items according to only one feature (e.g., a theme vowel), in situations where only one feature was relevant in said classification. However, the facts presented in the preceding chapters demonstrate that this is not the case in Latin or Romance verb conjugation, and that similarity must be judged using multiple features simultaneously. The proposal of subclasses necessarily implies that there is a hierarchy in which some features are more essential in classification than others. If it is true that no one feature is dominant in categorization, then there is no hierarchy. If there is no hierarchy, then there is no evidence for the establishment of subclasses.

In addition to the fact that positing subclasses branching from a single node is not necessarily warranted by the evidence, doing so makes some incorrect predictions. Once, say, the Latin conjugational class system makes its first split (traditionally into four or five classes based on the theme vowel), there is the implication that there is nothing more significant that certain verbs under one node share with certain verbs
under another node. However, this is not the case, as verbs of different traditional conjugational classes frequently do share other features, e.g., in Latin, verbs with different present stem types occasionally have the same perfect stem type. In Chapter 4, evidence from Balkan Romance shows that speakers must make connections that are inconsistent with a branching model (also shown in Juilland & Edwards 1971). Possibly the most convincing example discussed in this dissertation has to do with is the presence of a “velar augment” in Catalan. While the traditional notion of conjugational class in Catalan, as in the rest of Romance, is based on the theme vowel. However, the group of verbs that show this velar augment crosses these traditional conjugational classes. If Catalan speakers primarily classify verbs based on the theme vowel, then the generalization that verbs that share the characteristic of showing the velar augment do not all have the same theme vowel is lost. If Catalan speakers primarily classify verbs on the basis if they show the velar or not (i.e., that the highest node splits first into [+velar] and [-velar] classes), then the generalizations that certain verbs have the same theme vowel (while showing different behavior with respect to the velar augment) is lost. These cases demonstrate the need for an alternative model, specifically one that can deal with the simultaneous recognition of multiple features.

7.1.2. Advantages of this system

The model proposed here proves advantageous as it judges similarity simultaneously based on a variety of features. This is not to say that all features are necessarily judged equally, as there are likely some features of similarity that speakers judge as more salient and thus, weigh more heavily, in verbal categorization. However, the fact remains that there is insufficient evidence that the theme vowel is judged as being more significant than any other features.
This model deals with the notion of “conjugational patterns”, a term purposely used to provide distance to the traditional notions of “conjugational class”. As laid out in Chapter 1, any difference\(^2\) between the ways in which two verbs are conjugated signals that these two verbs follow two distinct “conjugational patterns”. Clearly, some of these patterns will be more similar than others, e.g., two certain patterns may be wildly different, while two other patterns might show a single difference. However, instead of grouping different patterns together in some sort of branching diagram, I propose that they are better of seen as forming a “constellation” (Janda & Joseph 1986). The patterns are organized with respect to the features they share, where these features include, but are not limited to, the theme vowel. Patterns that share more features are judged to be “closer together”, while patterns that share few features are judged as to be “further apart”. While all patterns do show some independence from one another, the borders between the different patterns are definitely permeable. Analogical change is more likely to occur with patterns that share more features (i.e., patterns that are more similar, and thus, are “closer together”), and certain patterns may have some sort of “attractive force”, which can be due to their size (patterns with a large number of followers are likely to attract more members), or due to other factors (some patterns are productive, even though this might not be the pattern with the largest number of followers), that incorporate new elements over time (either via language internal change or via language contact). However, this does not mean that changes that seem to move in the opposite direction than expected are not possible. In a sense, anything is possible, though some changes are typologically more frequent and have a greater likelihood to occur. As discussed in the previous chapters, analogical change, while often (and stereotypically) seen as a “simplifying” force can also move an the opposite

\(^2\) Of course, the only exceptions are differences that are attributable to the predictable phonological processes (i.e., processes that are phonetically motivated and operate in an environment without reference to morphological conditioning) of the language in question.
direction one would expect (e.g., the spread of 1SG -m in West and South Slavic, the fragmentation of the traditional 1st conjugation in Alguerese Catalan, etc.). For more on these and related issues in analogical change, see Wanner (2006).

This approach is definitely advantageous in describing the phenomena presented in the preceding chapters. For example, recall the traditional picture of the Daco–Romanian, where the 1st (a-stems) and 4th (i-stem) conjugations, are normally split into two subgroups depending on the presence of a stem extension ([EDZ] in a-stems, [ESC] in i-stems). The following diagram\(^3\) is indicative of this position:

![Diagram](image)

\textbf{Figure 7.1: Traditional view of DR 1st & 4th conjugations}

Recall that there is variation between the two subgroups of the 1st conjugation, between the two subgroups of the 4th conjugation, as well as between the subclasses of the 1st and 4th conjugation that show the stem extension. If we assume that, in a branching diagram, analogical change would frequent between two branches from a single node, then the variation between i-stem/[+ESC] and i-stem/[−ESC] types would be expected\(^4\) (they branch from the same node), and variation between a-stem/[+EDZ] and a-stem/[−EDZ] types would be expected as they too branch from the same node. However, the

---

\(^3\) For simplicity’s sake, this diagram only shows the relevant classes for the phenomena discussed here. For a more complete picture of the traditional views of Daco–Romanian conjugational classes, see Chapter 4.

\(^4\) “expected” probably is not the best choice of word here. Perhaps “unremarkable” or “unsurprising”. While it is rather tempting to “predict” or “expect” what will happen in analogical change, I find that so much of analogical change is unpredictable and unexpected that claiming to have this ability is dangerous.

217
common variation between a-stem/[+EDZ] and i-stem/[+ESC] types (e.g., as seen in the adaptation of modern loanwords, e.g., bloghez vs. bloghesc from Eng. blog) would not be expected under this tree model, because as soon as the highest node branches out into different conjugational classes based on the theme vowel, it makes the assumption that there is nothing more significant that certain “subclasses” of different classes share. Alternatively, under the view proposed here, these “subclasses” are considered independent “patterns” that are organized given the features they share. Under this model, the situation can be represented as shown below:

![Figure 7.2: Alternative view, DR a-stems & i-stems](image)

The two circles on the left represent patterns have the same specification for the thematic vowel, as do the two circles on the right. The two circles in the center share the feature of having a stem extension. The arrows represent common pathways of change. The above diagram is highly simplified, and it makes the untrue assumption that, e.g., all a-stem/[+EDZ] verbs conjugate in precisely the same way. When all of the potential features are taken into account, and even more complicated picture emerges.
Similar diagrams can be constructed with the Catalan data discussed in Chapter 5. It was shown that the verbs that show the velar augment are found across the traditional conjugational class boundaries. This view of conjugational patterns allows for there to be categorization based on both the theme vowel and on the presence/absence of the velar (among other factors), while not calling one feature “dominant” over the others.

Under some accounts of Latin/Romance verb morphology, irregular verbs are seen as almost “outside” of the system. In this model, irregular verbs, are in principle, no different from verbs normally considered “regular”. All patterns, are in some respect, equal, as they all exist. The patterns seen in irregular verbs do share certain characteristics with verbs seen as regular, so in the constellation of conjugational patterns, typically irregular verbs are just peripheral. For example, recall the example presented in Chapter 1 involving the traditional classification of the Spanish verbs tener and venir. There is a similar situation in Italian with the verbs tenere ‘hold–INF’ and venire ‘come–INF’. These two verbs have different theme vowels, as seen in the infinitive, the 2PL.PRES.IND. (tenete vs. venite), and the imperfect (e.g., tenevo ‘have–1SG.IMP.IND’ vs. venivo ‘come–1SG.IMP.IND’). As these verbs show different theme vowels, they are traditionally considered to be irregular verbs of different conjugations.

![Diagram](image)

Figure 7.3: Traditional view, It. tenere, venire, & verbs in –ere/–ire
However, this classification misses the generalization that these two verbs share several features to the exclusion of the other verbs in the conjugations in which they are traditionally classified. For example, both show the velar augment (e.g., *tengo* ‘hold–1SG.PRES.IND’, *vengo* ‘come–1SG.PRES.IND’), monophthong–diphthong alternation (e.g., *tiene* ‘hold–3SG.PRES.IND’, *viene* ‘come–3SG.PRES.IND’), similar patterns in the past tense (e.g., *tenni* ‘hold–1SG.PERF.IND’, *venni* ‘come–1SG.PERF.IND’), similar patterns in the future (e.g., *terrò* ‘hold–1SG.FUT’, *verrò* ‘come–1SG.FUT’), etc. Under the model proposed here, these similarities would definitely be taken into consideration. See the following diagram:

![Diagram](image)

Figure 7.4: Alternative view, It. *tenere*, *venire*, & verbs in –*ere* / –*ire*

Again, this diagram is highly simplified, as it makes the assumption that all Italian verbs in –*ere*, as well as –*ire*, respectively conjugate in precisely the same way, and this is untrue. This diagram only attempts to show the similarities that these two verbs have, along with the similarities they have with the “conjugations” with which they share a thematic vowel.

### 7.1.3. Applying this model elsewhere / the nature of language

While this dissertation has exclusively focused on Latin and Romance verbal morphology (along with the brief Sanskrit excursus in the first chapter), this is not to say that the
model proposed is only applicable to Romance verbal morphology. Similar issues are seen in other verbal systems.

For example, in Macedonian verbal morphology, there are a few conjugational patterns that verbs follow in the present tense, along with a few past (aorist) tense patterns. There is some level of predictability of the past tense pattern from the present (along with the present pattern from the past), though it is quite limited. Just taking the conjugation (via suffixation) of the present and aorist and linkages between the two one arrives at approximately eleven conjugation types (Friedman 2003). However, this becomes much more complicated when also taking aspectual morphology into account. The Macedonian present tense is of imperfective aspect, while the aorist is of perfective aspect. Perfective verbs are normally derived from imperfective verbs via prefixation (e.g., čita ‘read–IMPV’ vs. pročita ‘read–PRFV). However, there is a variety of prefixes used (e.g., pro–, do–, iz–, za–, etc.), and the specific one used in each case is not predictable. Thus, when taking these prefixes into account (along with the other linkages between imperfective verbs and their perfective counterpart), the picture is seriously fragmented. Situations such as this are not unusual whatsoever across the verbal systems of the Indo-European languages. A similar case is seen in Albanian. While some traditional grammars try to fit narrow the system into a few verb classes, there is really a wide range of patterns, e.g., Bega & Bega (2007) list 101 distinct verbal patterns.

The above examples are only concerned with the verb morphology of Indo-European languages. However, this is also seen in noun morphology of Indo-European languages that have maintained a case system. Even in languages that have lost the

---

5 For example, some perfective verbs are formed via suffixation of an imperfective verb, e.g., sluša ‘hear–IMPV’, slušne ‘hear–PRFV’, some imperfectives are derived from perfectives, e.g., kažuva ‘tell–IMPV’, kaže ‘tell–PRFV’, as well as cases where the imperfective and perfective are from completely different roots, e.g., gleda ‘see–IMPV’, vidi ‘see–PRFV’.

6 Dozon (1871) explains that there are attempts to describe Albanian as having two–conjugational systems, but it is quite problematic.
case system completely (e.g., Balkan Slavic, all of Romance except for Balkan Romance) there are similar trends in noun morphology. Think of English, for instance, a language that is considered to have limited morphology, and then think of all the patterns there are for forming plurals in English. There is one pattern that would be considered the “core” pattern, but the margins show a variety of patterns that cover just a few forms each. This is not even an Indo-European issue or a morphological issue. Similar cases are found in other languages, perhaps within or perhaps outside the domain of morphology. Instead of being a feature of Romance or Indo-European, this is a feature of language.

7.2. ON GENERALIZATIONS

7.2.1. The existence of generalizations

As described in Chapter 1, the model proposed here requires that all differences, regardless how “insignificant” they may look on the surface, between the ways in which two verbs are conjugated, show that these two verbs follow distinct patterns. Since verbs in any of the languages discussed here follow more than a few patterns, this model clearly assumes much of the information concerning verb conjugation is lexically-specific. That is, people have to have a whole lot memorized. What if this is taken even further? If so much has to be in the lexicon (i.e., a situation in which there are many patterns, some of which have hundreds of followers, while others have but one), then why not propose that everything is stored in the lexicon, i.e., a system in which there are no patterns whatsoever? Speakers have to already memorize many facts about each of these verbs, e.g., their phonetics, their semantics, any syntactic properties, etc. Why not just add the pattern in which they are conjugated to this list? The truth is that there is evidence that speakers categorize verbs, and are thus making generalizations. This is something that speakers are actively doing.
Sometimes it is difficult to observe such generalizations in a single synchronic state of a language. However, observing diachronic developments lets us observe this. As Kiparsky (1968) puts it, "What we really need is a window on the form of linguistic competence that is not obscured by factors like performance, about which next to nothing is known. In linguistic change we have precisely such a window". While looking at say, a specific dialect of Catalan or Daco–Romanian, it might be difficult to see what is going on, but when looking at the development of these varieties over time, it definitely gives us a window into what speakers are doing. They are clearly generalizing across patterns.

The word “generalization” has a bit of a connotation as being something involving simplification. However, as described in the previous section, it is not the case that all generalizations go in the direction of simplicity. There are situations where simplification is definitely the result, as in the case of stem allomorphy in the Latin verb POSSÔ being eliminated in the evolution to Spanish and Daco–Romanian. However, generalizations, even those seen as “local” can have an overwhelming affect of complication on verbal systems. For example, in Chapter 1, several examples were given of situations where an originally purely phonological rule became morphological. This often occurs in cases the where application of these phonological rules being extended beyond their original (purely phonological, and phonetically–motivated) environment. What is this extension? It is a generalization. A simple generalizations (to even one single form) in situations like this can have enormous complicating consequences.

It is given that humans generalize. Locally and based on similarity (the unknown is perceived in terms of the known). This is how we deal with everyday life. If humans did not generalize than every experience encountered each day would be a totally foreign experience.
7.2.2. The process of generalizing

As discussed above, there is ample evidence for the humans possess the ability to generalize. The process of generalizing is essentially the process of making a connection or fitting items into categories (i.e., grouping things together). However, how are these connections made?

Generalizations are made on the basis of similarity. People categorize similar items together. However, a deceptively tricky question has to do with what it actually means for two things to be similar? This seems like a ridiculously simple question, but it really is not.

I believe, that when it comes down to it, the task of judging similarity between objects involves the simultaneous evaluation of multiple features. This is not to say that one shared feature could result in two things appearing similar. For example, two people could be said to look “similar” if, say, they had similar-looking eyes. However, overall similarity between the appearance of people would be judged on more than this one feature.

This view is reminiscent of Wittgenstein’s “family-resemblance model” which he describes in his *Philosophical Investigations*, section 66. To illustrate this, he relates how the category of “games” is not defined by a single feature.

Consider …“games”. I mean board-games, card-games, ball-games, Olympic games, and so on. What is common to them all?... If you look at them you will not see something that is common to all, but similarities, relationships, and a whole series of them at that.... Look for example at board games, with their multifarious relationships. Now pass to card-games; here you find many correspondences with the first group, but many common features drop out, and others appear. When we pass next to ball-games, much that is common is retained, but much is lost. – Are they all amusing? Compare chess with noughts and crosses⁷. Or is there always winning and losing; but when a child throws his ball at the wall and catches it again, this feature has disappeared. Look at the

⁷ Also known as tic-tac-toe.
parts played by skill and luck; and at the difference between skill in chess and skill in tennis. Think now of games like ring-a-ring-a-roses; here is the element of amusement, but how many other characteristic features have disappeared! And we can go through the many, many other groups of games in the same way; can see how similarities crop up and disappear.

And the result of this examination is: we see a complicated network of similarities overlapping and cross–crossing: sometimes overall similarities, sometimes similarities of detail. (Wittgenstein 1967:31–32)

Just as the internal structure of the category we call “games” would be rather complex, the same can be seen in the structure of category we call “verbs” in terms of the conjugational characteristics they show. This “complicated network”, though represented a bit differently, is seen in the model proposed here (e.g., see Figures 7.2. and 7.4.). For example, looking at “games” again, we could say that the baseball and basketball are similar in that they are played on teams. We could also say that baseball and golf are similar in that they both involving striking a ball with some sort of instrument. However, just because baseball and golf are similar and baseball and basketball are similar, it does not entail that baseball and basketball are all that similar (except that they would both be considered “sports” or “games”, etc.). The same can be seen with verbs. In the previous section, the of Balkan Romance stem extensions was brought up again. Just because a-stem[+EDZ] and a-stem[−EDZ] classes are similar, and a-stem/[+EDZ] and i-stem/[+ESC] verbs are similar, it does not entail that a-stem/[−EDZ] and i-stem[+ESC] verbs are judged to be very similar.

How do we know that speakers make generalizations based on similarity? There is evidence in basically every aspect of human life, as Hahn & Ramsear (2001) put it, similarity and categorization “form the foundation of most intelligent human behavior”.

How does this apply to language? We can see the role of similarity in analogical change. Patterns that have more features in common are more likely to show variation (or have

---

8 A better example: person can look like their father and their mother (in terms of different features) simultaneously, while their father and mother look nothing alike.
members change between them) over time. For example, see the major trends in Daco–Romanian patterns involving stem extensions described in Chapter 4, as well as in section 7.1.2., above (and diagrammed in Figure 7.2). Various other cases of this are discussed in this dissertation.

7.2.3. On the scope of generalizations

Holyoak & Thagard (2005) write that analogy essentially entails “trying to understand the novel challenge in terms of what is already known, even if making the connection requires a mental leap.” Cases of analogical change (e.g., those described in Chapter 1) demonstrate that these “leaps”, are apparently not broad. Most of these cases of change are not cases of wide–spread generalizations, but rather generalizations of one form relating to one other form, or perhaps a small group of other forms. In cases where it looks like a wide–spread generalization has taken place, this is often reflective of a situation where many local generalizations have accumulated over time. For instance, the formation of the Balkan Romance verbs that show one of the stem extensions discussed in Chapter was not something that was done overwhelmingly at once. The group of verbs that show a stem extension grew one member at a time, in many cases accumulating verbs as they entered the language from external sources. The building (or in some cases, the present “dismantling”) of this collection of verbs is the result of “local generalizations”. The same can be seen in the spread of the velar augment in Catalan verbs. Once the spread started, it spread in different directions in different geographical areas and for different speakers. This spread occurred one form at a time, and thus, is the result of local generalizations. The whole point is that the history of verbal conjugational patterns and verbal categorization in Latin and Romance is marked by these local generalizations. Following the words of Kiparsky (1968), observing language change can give us insight into the synchronic state of language.
Thus, if the generalizations that led to the present situation were local in nature, why would we expect synchronic generalizations to be any different?

The data presented in the preceding chapters has shown that the generalizations that speakers of Romance languages make synchronically about their individual systems of verb classification and verb categorization are quite limited in scope. If we take the Gross Detail Method seriously, then the number of patterns exhibited by any of these languages is quite high. This is not to say that there are not generalizations that cover many forms. It is true that in any Romance language there will be patterns that cover hundreds of verbs. However, the majority of actual occurring patterns will only have a few (or potentially only one) followers.

### 7.3. On the Mind

It seems that linguists sometimes seem to impose limits on what speakers can and cannot do. Some solutions are rejected in that they are not “economical” enough or that they are missing a generalization (even if this generalization might be at the expense of the actual facts of the language). The truth is, though, that the human mind is capable of astounding feats. For example, the Mahābhārata, a Sanskrit epic longer than the Iliad and the Odyssey combined, was passed down orally for hundreds of years. Though it is an unquestionably impressive feat, one can potentially write this off as memorization. However, what about everyday life? How many faces are we capable of recognizing? How many intricate points of plots of books or movies are we able to recall, even though we have not read or seen them in years? The mind is capable of storing an immense amount of information.

Language is no different in these respects. We are capable of learning thousands upon thousands of words. We know how to pronounce them, what they mean, how to use them, when to use them, under what social situations their use is appropriate, and
so on. There is no legitimate reason to believe that language should operate so much more “economically” than other aspects of life (or that some aspects of language should operate more economically than others).

The question that must be brought concerns how speakers could possibly learn all these details. Upon observing the behavior of speakers, it is impossible to doubt the fact that they can. Any analysis of Romance verb morphology has to ultimately come to terms with the fact that native speakers of a language do know all of these forms. While the model advocated in this dissertation may look more complicated than necessary on the surface, in this respect it really is not any more complicated than any other model. However, the advantage the model discussed here is that is actually takes the fact that speakers do know all of these forms very seriously, and, at least in my opinion, provides a more psychologically realistic view than the traditional accounts.

Linguists frequently believe that speakers prefer simpler solutions. Definitely, there are cases where this looks like it is true. There are numerous cases of paradigms being leveled or morphological categories being lost. However, speakers do not always prefer the simpler solution. As discussed in Chapter 1, German has over sixty umlaut rules, Sanskrit had dozens of reduplication rules (Janda & Joseph 1986,1989), and as shown in Chapter 2, there are dozens of ways in which Latin verbs were conjugated once all of the facts are sorted out. Is all of this complication really “necessary”? No, this could have been much simpler, e.g., German could just have one umlaut rule, Sanskrit could have had just one reduplication rule, and Latin could have had just a single pattern following which all verbs were conjugated. However, instead of coming up with a simpler answer, speakers of these languages took a more complex path.

If speakers of Romance languages were really concerned with simplicity, there would be no velar augments in Catalan, the stem extensions [ESC] and [EDZ] would have been eliminated from Balkan Romance, and the whole notion of present – perfect –
participial stem linkages would not be necessary, as all conjugated forms would be formed from the same stem. All of these patterns would be wiped out in favor of something more consistent, something less irregular. However, this is just not how language works and this just is not how the human mind works. Linguistic proposals that favor simplicity at the expense of the facts in order to propose a more economical system should step back and realize that such proposals go against the nature of both language and the human mind. Language is not simple and the human mind is more than capable of dealing with this fact.
REFERENCES


COSTANZO, ANGELO. 2008. On the origin, expansion and productivity of a Romanian conjugational (sub-) class. Columbus: The Ohio State University. Unpublished manuscript.


Fabra, Pompeu. 1898. *Contribució a la gramàtica de la llengua catalana.* Barcelona: L’Avenç.


238


239


PÉREZ SALDANYA, MANUEL; MANUEL SIFRE GÓMEZ; and JÚLIA TODELLI CERVERA. 2004. *Morfologia catalana*. Barcelona: Editorial UOC.


