



A Morphological Typology of Non-Root Alternations: Invasion, Suppletion, and Allomorphy

Citation

Kühlert, Niels Torben. 2023. A Morphological Typology of Non-Root Alternations: Invasion, Suppletion, and Allomorphy. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37378053>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

HARVARD
Kenneth C. Griffin



GRADUATE SCHOOL
OF ARTS AND SCIENCES

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Linguistics
have examined a dissertation entitled

A Morphological Typology of Non-Root Alternations: Invasion, Suppletion, and
Allomorphy

presented by Niels Torben Köhlert

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature Jay H. Jasanoff

Typed name: Prof. Jay Jasanoff, co-chair

Signature Jeremy Rau

Typed name: Prof. Jeremy Rau, co-chair

Signature Jonathan Bobaljik

Typed name: Prof. Jonathan Bobaljik

Date: July 13, 2023

A Morphological Typology of
Non-Root Alternations:
Invasion, Suppletion, and Allomorphy

a dissertation presented by
Niels Torben Kühlert
to the Department of Linguistics

in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in the subject of Linguistics

Harvard University
Cambridge, Massachusetts

July 2023

©2023 – NIELS TORBEN KÜHLERT

ALL RIGHTS RESERVED.

A Morphological Typology of Non-Root Alternations::

Invasion, Suppletion, and Allomorphy

ABSTRACT

This dissertation investigates analogy, with the goal being to establish an empirical and theoretical basis for determining which factors influence analogical changes. It centers on an understudied set of phenomena known as *non-root intra-paradigmatic alternations*, with a specific emphasis given to those alternations whose distributions do not conform to natural classes. Overall, this dissertation aims to (i) demonstrate the value of studying non-root intra-paradigmatic alternations, (ii) set a basis for constraining analogical change by examining these irregular alternations, and (iii) show how different factors affect the directionality of paradigmatic leveling.

By examining a wide range of language data, this dissertation provides novel contributions pertaining to non-root alternations and their interaction with analogical change. Firstly, it presents a full typology of non-root alternations, based on the different origins from which they arise. This dissertation argues that two types of morphological phenomena, affixal and synthetic-analytic alternations, show similar characteristics both in terms of their diachronic development and synchronic distribution. Treated together, close examination of these distributions provides insights into the motivations for analogical change and the synchronic structure of grammar.

Additionally, this dissertation demonstrates that irregular alternations are governed by abstract generalizations which should be extendable to new lexical items. By providing a detailed diachronic account for the development of periphrasis in the Latin verb *nōlō* ‘I do not want’, it shows that synthetic-analytic alternations can be extended analogically on the basis of abstract generalizable patterns. Furthermore, this dissertation investigates the interaction of irregular distributions with analogical change, focusing on how paradigmatic leveling has affected the distribution of the *-sc-* suffix in several varieties of Romance. In doing so, it posits several biases involving frequency and a desire to adhere to natural morphosyntactic classes which provide a basis for explaining the attested directionality of change.

Meinen Eltern

ACKNOWLEDGEMENTS

This dissertation is the culmination of a journey marked by invaluable contributions and unwavering support from many different people. I extend my heartfelt gratitude to each of you. I would like to begin by giving my deepest thanks to my advisors. To Jay, I am grateful for your advice, insightful feedback, and direction in approaching historical problems. To Jeremy, I am incredibly thankful for your continuous encouragement as well as for the many hours of engaging discussion. I would also like to thank Jonathan for your guidance, critical perspectives, and our thought-provoking conversations over the years.

A special appreciation goes out to my colleagues and peers whose insights have proved indispensable in the creation of this dissertation. In particular, I would like to thank Michael Weiss, Michael Flier, Adam Albright, Kevin Ryan, Aljoša Milenković, Stan Zompì, and Christos Christopoulos for our discussions regarding both empirical and theoretical aspects of this work. Additionally, I am grateful (in no particular order) to Shannon Bryant, Natasha Thalluri, Agnes Ruyue Bi, Peter Grishin, Jim Huang, Yvette Yi-Chi Wu, Joshua Martin, Wei-Fang Hsieh, and Zachary Rothstein-Dowden for their support and good times over the years. A special mention goes out to Ian Kirby, who has been with me since the beginning of this journey.

To Jonathan Kenkel, Robert Wild, and Siddhartha Jayanti, our friendship will always remain strong. Above all, I would like to thank my family, especially my parents, my sister, and of course Tamisha for your unwavering support and enduring patience through all the challenges. Without your help, I would not be where I am today.

Table of Contents

TITLE PAGE	1
COPYRIGHT	ii
ABSTRACT	iii
ACKNOWLEDGMENTS	v
ABBREVIATIONS	viii
LIST OF TABLES	ix
LIST OF FIGURES	xiii
1 INTRODUCTION	1
1.1 Analogy	1
1.2 Non-root intra-paradigmatic alternations	3
1.3 Overview of different types of analogy	5
1.4 Previous approaches to morphological learning	10
1.5 Main claims	19
1.6 Roadmap	23
2 NON-ROOT INTRA-PARADIGMATIC ALTERNATIONS	25
2.1 Defining non-root intra-paradigmatic alternations	26
2.2 Diachronic typology of alternations	37
2.3 Types of non-root alternations	40
2.4 Morphomic distributions	74
2.5 Alternations and synchrony	84
2.6 Conclusion	87
3 SYNTHETIC-ANALYTIC ALTERNATIONS	88
3.1 Introduction	88
3.2 Establishing the facts	90

3.3	Diachronic Explanation	129
3.4	Latin morpheme: synchronic perspective	148
3.5	Conclusion	186
4	IRREGULAR DISTRIBUTIONS AND LEVELING	187
4.1	Introduction	187
4.2	Background	195
4.3	Proposal	218
4.4	Extension of <i>-sc-</i> based on person + number	228
4.5	Leveling of the subjunctive	246
4.6	Conclusion	264
5	CONCLUSION	270
5.1	Overview	270
5.2	Contributions and insights	271
5.3	Directions for further work	274
	APPENDIX A LATIN ADDENDUM	279
A.1	Paradigms	279
A.2	Counts of <i>nōlō</i> and <i>mālō</i> in Plautus	283
A.3	Secondary Factors: usage and frequency	287
A.4	Vocabulary Items	292
A.5	Stem Functions	295
A.6	Extension to <i>nōlō</i>	298
	APPENDIX B ROMANCE ADDENDUM	300
B.1	Background on Latin verbal system	300
B.2	Romance verbal system	304
B.3	Non-patterned distributions	308
	REFERENCES	314

Abbreviations and glosses

ACT	active	PL.	Plautus
ATT.	Attic Greek	PLPF	pluperfect
AUTH	author	PERF	perfect
CAECIL.	Caecilius	PRES	present
CONJ	conjunction	PRET	preterite
IMPF	imperfect	PROM.	Proto-Romance
LAT.	Latin	PTPL	participle
NAEV.	Naevius	RUSS.	Russian
NON.	Nonius	SMP PST	simple past
OCS	Old Church Slavonic	SUBJ	subjunctive
OHG	Old High German	TER.	Terence
OLAT.	Old Latin	VED.	Vedic Sanskrit
PART	participant		

Plays of Plautus:

AMPH.	Amphitruo	MIL.	Miles Gloriosus
ASIN.	Asinaria	MOST.	Mostellaria
AUL.	Aulularia	PERS.	Persa
BACCH.	Bacchides	POEN.	Poenulus
CAP.	Captivi	PSEUD.	Pseudolus
CAS.	Casina	RUD.	Rudens
CIST.	Cistellaria	STICH.	Stichus
CURC.	Curculio	TRIN.	Trinummus
EPID.	Epidicus	TRUC.	Truculentus
MEN.	Menaechmi	VID.	Vidularia
MERC.	Mercator		

List of Tables

1.1	PRES IND of <i>andàre</i> ‘to go’ and <i>finìre</i> ‘to finish’ (Italian)	4
2.1	PRES IND of <i>lucra</i> ‘to work’ (Romanian)	27
2.2	<i>medvežónok</i> ‘bear cub’ (Russian)	27
2.3	PRES IND of <i>nōlle</i> ‘to not want’ and <i>velle</i> ‘to want’ (Latin)	28
2.4	<i>žena</i> ‘wife’ in Prizren-Timok (Svinitsa)	29
2.5	Extended paradigm of <i>lucra</i> ‘to work’ (Romanian)	32
2.6	<i>čelovek</i> ‘person’ (Russian)	36
2.7	Root Suppletion Alternations for ‘to go’ in Romance	39
2.8	IMPF ACT IND of <i>amāre</i> ‘to love’ (Latin)	41
2.9	IMPF PASS IND of <i>amāre</i> ‘to love’ (Latin)	42
2.10	IMPF IND of <i>zvati</i> ‘to call’ (Old Church Slavonic)	43
2.11	PRET IND of <i>sruk-</i> ‘die’ (Tocharian B)	43
2.12	δότηρ <i>dótōr</i> ‘giver’ (Ancient Greek)	45
2.13	PRES IND of <i>finir</i> ‘to finish’ (French)	46
2.14	PRES IND of <i>*finiscō</i> ‘finish’ (Proto-Romance)	47
2.15	PRES IND of <i>lucra</i> ‘to work’ (Romanian)	48
2.16	PRES IND of <i>ne byti</i> ‘to not be’ (Old Church Slavonic)	49
2.17	άνθρωπος <i>ánt^hrōpos</i> ‘man’ (Attic)	49
2.18	Potential decomposition of the thematic declension (Attic)	50
2.19	$\bar{\alpha}$ -declension (Attic)	55
2.20	Development of <i>o</i> -stem plurals (Greek)	60
2.21	Development of the $\bar{a}$ -stem plurals (Greek)	60
2.22	FUT of <i>ferō</i> ‘I bear’ (Latin)	62
2.23	PRES SUBJ of <i>ferō</i> ‘to bear’ (Latin)	63
2.24	IMP of <i>věděti</i> ‘to lead’ (Old Church Slavonic)	65
2.25	<i>medvežónok</i> ‘bear cub’ (Russian)	68
2.26	<i>seljanín</i> ‘villager’ (Russian)	70
2.27	N-PATTERN in Italian, Catalan, and Romanian	70
2.28	PRES IND of <i>ferō</i> ‘I bear’ (Latin)	71
2.29	PRES IND of <i>ferō</i> ‘I bear’ and <i>gerō</i> ‘I bear’ (Latin)	72
2.30	PRES IND of <i>andàre</i> ‘to go’ and <i>finìre</i> ‘to finish’ (Italian)	77

2.31	Morphomic template (Italian)	77
3.1	PRES IND of <i>volō</i> , <i>nōlō</i> , and <i>mālō</i> (Latin)	90
3.2	PRES IND of <i>colō</i> ‘I cultivate’ and <i>volō</i> ‘I want’ (Latin)	94
3.3	PRES INF of <i>volō</i> and <i>colō</i> (Latin)	95
3.4	IMPF SUBJ for <i>volō</i> and <i>colō</i> (Latin)	95
3.5	PRES IND of <i>nōlō</i> ‘I not want’ (Latin)	96
3.6	PRES IND of <i>mālō</i> ‘I prefer’ (Latin)	97
3.7	PRES INF of <i>nōlō</i> and <i>mālō</i> (Latin)	97
3.8	IMPF SUBJ of <i>nōlō</i> and <i>mālō</i> (Latin)	97
3.9	PRES IND of <i>volō</i> (Old Latin/Classical Latin)	99
3.10	PRES IND of <i>nōlō</i> in Plautus (Counts)	100
3.11	<i>nōlō</i> (Old Latin/Classical Latin)	102
3.12	PRES IND of <i>mālō</i> in Plautus (Counts)	102
3.13	SYNTHETIC/ANALYTIC forms in the PRES SUBJ of <i>nōlō</i> (Old Latin)	105
3.14	SYNTHETIC/SEMI-ANALYTIC forms in the PRES SUBJ of <i>mālō</i> (Old Latin)	105
3.15	PRES SUBJ of <i>volō</i> in Plautus (Counts)	106
3.16	IMPF SUBJ of <i>nōlō</i> and <i>mālō</i> in Plautus (Counts)	106
3.17	IMPF SUBJ of <i>volō</i> in Plautus (Counts)	107
3.18	Preference for forms of the PRES and IMPF SUBJ in Plautus	108
3.19	2SG and 3SG singular of <i>nōlō</i> in Plautus (Counts)	111
3.20	PRES IND of <i>ferō</i> and <i>gerō</i> (Latin)	114
3.21	PRES INF of <i>ferō</i> and <i>gerō</i> (Latin)	114
3.22	IMPF SUBJ of <i>ferō</i> and <i>gerō</i> (Latin)	114
3.23	PRES IND of <i>sum</i> , <i>edō</i> , and <i>possum</i> (Latin)	115
3.24	PRES INF of <i>sum</i> and <i>edō</i> (Latin)	116
3.25	IMPF SUBJ of <i>sum</i> and <i>edō</i> (Latin)	116
3.26	Patterned athematicity in the PRES IND	117
3.27	Pre-Latin paradigm for <i>volō</i>	120
3.28	Pre-univerbated paradigm for <i>*ne uolō</i>	130
3.29	Final distribution of <i>nōlō</i> (Classical Latin)	138
3.30	Summary of athematic behavior (Latin)	151
3.31	IMPF and FUT IND of SUM (Latin)	152
3.32	PRES SUBJ of <i>FERō</i> , <i>VOLō</i> , <i>SUM</i> , <i>EDō</i> (Latin)	156
3.33	Irregular PERF forms (Italian)	163
3.34	<i>parō</i> ‘I prepare’ (Latin)	173
4.1	PRES IND/SUBJ of <i>finire</i> ‘to finish’ (Italian)	189
4.2	Italian <i>finire</i> , Catalan <i>patir</i> , Romanian <i>a păți</i>	189
4.3	N-PATTERN (Romance)	200
4.4	Vowel alternations in stems (Romance)	202

4.5	Stress in Latin conjugation	202
4.6	‘to lose’ in Latin, Italian, Spanish, and Romanian	203
4.7	Suppletive alternations for ‘to go’ (Romance)	205
4.8	PRES IND/SUBJ of <i>uscìre</i> ‘to go out’ (Italian)	205
4.9	Reflexes of <i>-idi-</i> (Romance)	207
4.10	N-PATTERN (Italian)	212
4.11	Forms of ANDÀRE ‘to go’ (Italian)	213
4.12	Forms of FINÌRE ‘to finish’ (Italian)	213
4.13	PRES IND/SUBJ of <i>polìr</i> ‘to clean’ (Mantuan)	229
4.14	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Gavardo)	229
4.15	PRES IND/SUBJ of <i>córar</i> ‘to run’ (Mantuan)	229
4.16	PRES IND/SUBJ of <i>ríder</i> ‘to laugh’ (Gavardo)	229
4.17	PRES IND/SUBJ of <i>finìi</i> ‘to finish’ (Cremonese)	230
4.18	PRES IND/SUBJ of <i>capi</i> ‘to understand’ (Bergamasque)	230
4.19	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Paves)	230
4.20	PRES IND/SUBJ of <i>parlàa</i> ‘to speak’ (Cremonese)	231
4.21	PRES IND/SUBJ of <i>crèd</i> ‘to believe’ (Bergamasque)	231
4.22	PRES IND/SUBJ of <i>andàa</i> ‘to go’ (Cremonese)	236
4.23	PRES IND/SUBJ of <i>andà</i> ‘to go’ (Milanese)	236
4.24	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Piedmontese)	238
4.25	PRES IND/SUBJ of <i>canté</i> ‘to sing’ (Piedmontese)	239
4.26	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Piedmontese)	240
4.27	PRES IND/SUBJ of <i>andé</i> ‘to go’ (Piedmontese)	240
4.28	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Piedmontese)	243
4.29	PRES IND/SUBJ of <i>capi</i> ‘to understand’ (Milanese)	247
4.30	PRES IND/SUBJ of <i>vedè</i> ‘to see’ (Milanese)	248
4.31	PRES IND/SUBJ of <i>patí</i> ‘to suffer’ (Ligurian)	248
4.32	PRES IND/SUBJ of <i>fin-</i> ‘to finish’ in (Upper-Engadinian)	248
4.33	PRES IND/SUBJ of <i>florír</i> ‘to flourish’ (Old Occitan)	249
4.34	PRES IND/SUBJ of <i>cöglie</i> ‘to pick up’ (Ligurian)	249
4.35	IMPERF IND and PRES SUBJ of <i>sentì</i> and <i>capi</i> (Milanese)	250
4.36	PRES IND/SUBJ of <i>cantár</i> ‘to sing’ (Old Occitan)	252
4.37	PRES IND/SUBJ of <i>florír</i> ‘to flourish’ (Old Occitan)	252
4.38	PRES IND/SUBJ of <i>decir</i> ‘to say’ (Spanish)	254
4.39	PRES IND/SUBJ of <i>valier</i> ‘to be worth’ (Old French)	254
4.40	PRES IND/SUBJ of <i>poder</i> ‘to be able’ (Portuguese)	256
4.41	PRES IND/SUBJ of <i>pudài</i> ‘to be able’ (Cascinagrossa)	256
4.42	PRES IND/SUBJ of <i>andé</i> ‘to go’ (Piedmontese)	259
4.43	PRES IND/SUBJ of <i>andà</i> ‘to go’ (Ligurian)	260
4.44	PRES IND/SUBJ of <i>anar</i> ‘to go’ (Old Occitan)	260
4.45	PRES IND/SUBJ of <i>ir</i> ‘to go’ (Upper-Engadinian)	260

4.46	FUT IND and PRES COND of <i>capi</i> ‘to understand’ (Milanese)	265
4.47	PRES COND of <i>finì</i> ‘to finish’ (Paves)	266
4.48	PRES IND/SUBJ of <i>finì</i> ‘to finish’ (Piedmontese)	267
A.1	Finite forms of <i>volō</i> ‘I want’ (Latin)	280
A.2	Finite forms of <i>nōlō</i> ‘I do not want’ (Latin)	281
A.3	Finite forms of <i>mālō</i> ‘I prefer’ (Latin)	282
A.4	Regular forms of PRES IND (<i>nōlō</i>)	283
A.5	Other negated forms in the PRES IND	283
A.6	IMPF IND of <i>nōlō</i>	284
A.7	IMPF SUBJ of <i>nōlō</i>	284
A.8	FUT of <i>nōlō</i>	284
A.9	PERF IND of <i>nōlō</i>	285
A.10	Other negated perfects of <i>nōlō</i> ‘I do not want’	285
A.11	Other negated forms of <i>nōlō</i> ‘I do not want’	285
A.12	Regular forms of PRES IND (<i>mālō</i>)	285
A.13	PRES IND of <i>volō</i> in Plautus (Counts)	289
A.14	PRES IND <i>nōlō</i> in Plautus (Counts)	289
A.15	PRES IND of <i>mālō</i> in Plautus (Counts)	289
A.16	‘Child-directed’ forms of <i>volere</i> ‘to want’ (Italian)	291
A.17	Child-uttered forms of <i>volere</i> ‘to want’ (Italian)	291
A.18	PRES SUBJ for <i>sum</i> ‘I am’ (Latin)	293
A.19	PRES SUBJ (Latin)	294
B.1	<i>infectum</i> and <i>perfectum</i> split (Latin)	301
B.2	Suppletive split for <i>ferō</i> ‘I bear’ (Latin)	301
B.3	<i>calēscō</i> ‘I become hot’ (Latin)	302
B.4	Overview of conjugation (Latin)	303
B.5	Categorical changes from Latin to Modern Romance	305
B.6	<i>padecer</i> ‘to suffer’ (Spanish)	309
B.7	<i>padecer</i> ‘to suffer’ (Portuguese)	309
B.8	Finite forms with <i>-iss</i> < <i>*-isc-</i> (French)	311
B.9	PRES IND of verbs with N-PATTERN (French)	312

List of Figures

2.1	Types of non-root alternations 1	40
2.2	Types of non-root alternations 2	74
B.1	Romance languages	308

1

Introduction

1.1 ANALOGY

Despite being one of the primary driving forces behind language change, analogy remains far less understood than its more familiar cousin, sound change. This is not to say that progress has not been made with regards to our understanding of how analogy occurs and what factors trigger and influence the directionality of change. In particular, several recent research programs have focused on how frequency and notions of morphological productivity intersect to produce analogical effects, whether it be through

measuring *hapax legomena* (Sandell, 2015) or through a learning algorithm like the Tolerance Principle (Yang, 2016, Kodner, 2022). However, there are still significant gaps when it comes to formalizing analogy and determining the constraints that limit when analogy can and cannot take place.

In what follows, I take the position that trying to constrain analogical change should be one of the main goals of those who study historical change. Even though many have cast doubt on the idea that constraining analogy is a worthwhile endeavor (e.g. Delbrück 1893, Paul 1886, Spina & Dressler 2011: 535),¹ I argue that despite the difficulty in doing so, trying to figure out the limitations of analogical change is crucial in furthering our knowledge of not only language structure but also linguistic change over time.

Firstly, establishing an empirical basis of what is and is not possible in analogical change would vastly improve the descriptive power of analyses which make reference to the mechanism of analogy. Those who study historical linguistics often reference a vague intuition of what is a reasonable analogy; eliciting exactly what this intuition entails is crucial in assessing the plausibility of historical hypotheses with a greater degree of rigor.

Secondly, we should be aiming not only to describe changes but also to account for why and how analogical changes occur from a general perspective. If we wish to build a solid intuition of what is reasonable, we need to know how analogy works. It is clear that analogy does not arise randomly; ever since its invocation by the Neo-grammarians, analogy has been known to follow certain principles (cf. Paul 1886). Given this, it is worth trying to understand in what ways analogy is both systematic and limited.

While I do not intend to solve all of these issues here, this dissertation aims to

¹See also the discussion in Fertig (2013: 102-103) for more specifics and other citations. Additionally, some authors have expressed doubt that diachronic change in general cannot be predicted and thus cannot be constrained (Hock 1991, Lass 1997, Croft 2000, *a.o.*).

set an empirical and theoretical basis for determining the different biases that influence learners into making analogical changes. By examining a previously understudied set of phenomena, *non-root intra-paradigmatic alternations*, I identify several biases which affect analogical leveling and proportional change. These biases include various usage-based factors including frequency as well as structural biases which revolve around syntactic and morphological coherence. Especially informative are *morphomic* distributions, idiosyncratic alternations which cannot be explained by a straightforward morphosyntactic analysis. I give specific focus to how these distributions intersect with analogical change, both in terms of extension and leveling.

Before going into more detail about the specific biases I propose, I discuss two key elements of this dissertation. The first are *non-root intra-paradigmatic alternations*, which are the main phenomena discussed in this dissertation. Secondly, I briefly review the different types of analogy and some of the relevant learning algorithms that have been proposed to account for analogical change.

1.2 NON-ROOT INTRA-PARADIGMATIC ALTERNATIONS

This section briefly describes the phenomena discussed in this dissertation, prefiguring the discussion in Section 2.1. Putting aside the notion of *intra-paradigmatic* for now, the term *non-root alternations* refers to alternations which are expressed by morphological marking which does not involve the root specifically. To illustrate this, compare the present indicative of two Italian verbs below:

	ANDÀRE ‘to go’	FINÌRE ‘to finish’
1SG	<i>vàd-o</i>	<i>finisc-o</i>
2SG	<i>và-i</i>	<i>finisc-i</i>
3SG	<i>và</i>	<i>finisc-e</i>
1PL	<i>and-iàmo</i>	<i>fin-iàmo</i>
2PL	<i>and-àte</i>	<i>fin-ite</i>
3PL	<i>vàn-no</i>	<i>finisc-ono</i>

Table 1.1: PRES IND of *andàre* ‘to go’ and *finire* ‘to finish’ in Italian

These two paradigms both contain morphological alternations affecting the stems of the verb. For *andare*, *and-* contrasts with *va(d)-*, while *finire* contrasts the stem *finisc-* with *fin-*.

Despite the fact that both paradigms alternate their stems, there is a crucial difference between the two verbs. For *andare*, the root of the verb suppletes completely; speakers of Italian must store two allomorphs for the root of the verb, *va(d)* and *and*. On the other hand, there is only one allomorph for *finire*, namely *fin*. The difference between *finisc* and *fin* concerns the extra segment *-isc-*, rather than the root. Thus, *andare* contrasts root allomorphs (root alternation), while *finire* contrasts allomorphs of the suffix *-isc-/∅* (non-root alternation).

There are many types of non-root alternations, which will be expanded upon in Chapter 2. The generalized principle to keep in mind is that all of these alternations are expressed in terms of morphology which does not contain the root. Thus, the root always remains the same. The notion of *intra-paradigmatic* refers to the fact that the alternation takes place in a limited part of a paradigm. In the forms in Table 1.1, the alternation is limited to the present indicative. Section 2.1.2 will further expand on and clarify this notion.

Although I expand on this point later, the forms in Table 1.1 crucially express a distributional coherence, which does not seem to be a coincidence. Despite being completely separate phenomena, suppletion and affixation pattern together in an idiosyncratic distribution. From a historical perspective, this relation hints at some sort of analogical comparison between these two phenomena. However, as we will see, it is not exactly clear how this relationship should be described in traditional terms. Section 2.4 expands on these types of distributions, known as *morphomes*. Much of this dissertation focuses on morphomes and how they intersect with analogy.

1.3 OVERVIEW OF DIFFERENT TYPES OF ANALOGY

Over the years, analogy has been classified into many different types.² This section does not aim to provide a novel classification of analogy. Nor does this section intend to be particularly controversial in its description. It simply presents a broad overview of the different types of analogy as will be relevant for further discussion later in this dissertation. That being said, the following sections discuss proportional change, non-proportional change, and leveling.³

1.3.1 PROPORTIONAL CHANGE

Proportional change broadly refers to analogical changes which make reference to a base by which a proportion can be made. Schematically, such changes are framed in terms of the following equation: $A : A' :: B : X$, $X = B'$. In other words, the relationship between a form A and a derived form, A' , can be applied to a separate form B , producing form

²For some of these, one may consult the standard textbooks describing historical linguistics and language change, e.g. [Hock & Joseph \(2009\)](#), [Crowley & Bowerman \(2010\)](#), [Campbell \(2013\)](#), *a.o.* See also the classifications in [Andersen \(1980\)](#), [Fertig \(2013\)](#). One may also wish to consult some of the foundational material on analogy, including [Paul \(1886\)](#), [Wheeler \(1887\)](#), among other authors like Hermann Osthoff and Karl Brugmann.

³This classification is loosely based on that proposed in [Fertig \(2013: 83\)](#) and [Hock & Joseph \(2009\)](#).

X , which is solved in the equation as B' . This type of proportional relationship can be exemplified by the classic example of English plurals (cf. Hock & Joseph 2009: 157): *stone* : *stone-s* :: *cow* : X , $X = \textit{cow-s}$. This type of schema has been traditionally referred to as a *four-part analogy*. Four-part analogy does not necessarily imply that it is the two lexemes in particular that are compared; rather, four-part analogy represents the productive process by which new forms are built on a set of already learned forms.⁴ In this sense, *cow* is compared not only to *stone* but all the nouns that form their plural with *-s*.

There are additional terms that have been used to categorize changes as proportional. I focus here on *analogical extension* and *backformation*. The former refers to when a form (such as an inflectional suffix) or a morphological pattern is applied to a grammatical form which it did not originally apply to. The application of the plural marker *-s* to *cow* in the four-part analogy illustrated above could be described in terms of an extension. Here, the suffix *-s* has been extended to *cows*, replacing the older plural form *kine*. Other examples include voicing alternations in stem-final consonants in English: *scarf-scarves* [f~v], replacing *scarfs* [f] (see Fertig 2013: 50 for this and more examples).

The term *backformation* refers to the process by which speakers (mis)parse the morphological composition of a construction, allowing for the creation of novel forms. A classic example is the English word *pea* which is a newly created singular form, following a misparsing of an earlier *pease* which was a mass noun (Hock & Joseph 2009: 160). Given that the end of *pease* was phonologically identical to the plural suffix *-s*, speakers reinterpreted the form as a plural, *backforming* a singular form without the *-s*.

It is important to note that the differences between four-part analogy, extension, and

⁴See the discussion of Paul's conception of proportional equations in Fertig (2013: 138-139).

back-formation really are rather blurry, as the latter two may often also be described in terms of a proportional equation with the structure $A : A' :: B : X, X = B'$, and as such are proportional. As we have seen, the application of *-s* to *cow* to produce *cows* is both an example of four-part analogy and analogical extension. Instead of assuming that each of these terms substantiates a different type of analogy, I will treat all of these as examples of a singular type of analogy, with the term *proportional analogy*.

1.3.2 NON-PROPORTIONAL CHANGE

Not all types of analogical change involve specific equations or proportions. There are several types that involve highly idiosyncratic changes, which cannot easily be described schematically, and are hence *non-proportional*. Among the types that have been described as such are *folk etymology*, *contamination*, and *blending*. As this dissertation does not focus on non-proportional analogy, I will only briefly describe these types.

Folk etymology broadly refers to when speakers misanalyze forms, usually based on phonological or semantic criteria. An example which illustrates this is the English word *bridegroom*, whose Old English predecessor was *brȳdguma* (Fertig 2013: 10),⁵ which itself was a compound of ‘bride’ (*brȳd*) + ‘man’ (*guma*). The insertion of the *-r-* in *guma* was ultimately due to speakers extrapolating a non-etymological connection with the unrelated word *groom*. The difference between *folk etymology* and *backformation* is that the latter involves the production of new forms based on a morphological misparsing, while the former involves a misanalysis of a specific form.

Contamination refers to the process by which semantically related forms may influence each others phonological form. These are often found in antonyms or in numerals. For example, the numeral nine in Slavic changed its initial consonant from an **n-* to

⁵cf. German *Bräutigam*.

a **d-* on the basis of the numeral for ten, which began with a **d-*. Thus, in Russian we find *dévjat* ‘nine’ and *désjat* ‘ten’, both beginning with *d-*. The *d-* in *dévjat* has been extended from *désjat*. Other Indo-European languages retain the *n-*: Lat. *novem* ‘nine’ vs *decem* ‘ten’.

Finally, *blending* refers to the concatenation of two words to produce an often humorous portmanteau form which is not neatly decomposable morphologically. Examples may include words like *brunch* < *breakfast* + *lunch* or *motel* < *motor* + *hotel*.

None of these changes may easily be explained with a specific proportion. For example, one could not describe the relationship between Russ. *dévjat* and *désjat* in terms of changes to *A*’ or *B*’. These analogies are highly irregular and represent some of the most idiosyncratic types of morphological change. For this reason, we should expect the factors that govern the creation of these forms to be different than those that govern proportional change. Thus, I put aside these changes in this dissertation.

1.3.3 LEVELING

The final type of analogy discussed here is *leveling*. This type refers to the process by which the number of allomorphs in a given paradigm is reduced.

Leveling is often exemplified by the replacement of stems within a paradigm. For example, the comparative and superlative of the adjective *old* were once *elder* and *eldest* respectively. The difference in vowel quality was once conditioned by umlaut, a regular phonological process; however, the alternation has since been morphologized. While *elder* and *eldest* were retained in specific use cases, they were generally replaced by *older* and *oldest*. Effectively, the stem *eld-* was replaced by the regular adjectival form *old-*. Countless other examples can be found in all sorts of nominal and verbal paradigms cross-linguistically.

Some classifications of the different types of analogy subsume leveling under proportional change. Indeed, leveling is accompanied by an amount of systematicity which truly separates it from non-proportional changes. Additionally, one can often make proportions that model the leveling, either entirely or at least partly. For example, one could explain the creation of *older* through a proportion with an adjective that did not have a stem alternation, e.g. *young* : *younger* :: *old* : *X*, *X* = *older*. In this sense, there might be a reason to categorize leveling as simply another type of proportional change.

Whether leveling falls under proportional change or should be classified as its own unique phenomenon is not just a terminological issue; rather, this question directly relates to how we analyze cases of leveling. If leveling is a type of proportional change, we would expect speakers to make direct comparisons between paradigms with alternations and paradigms without and through this comparison motivate the leveling. If leveling is not proportional, an external comparison is not needed, and forms within a paradigm only need to reference one another in order to motivate leveling. There has been considerable debate in the literature about whether leveling should be considered proportional based on these two aforementioned positions.⁶

I take the position that a simple proportional analysis of leveling fails to account for the directionality present in examples of paradigmatic leveling from Romance.⁷ In this sense, this dissertation contributes to the position that leveling is a unique type of analogy that has its own properties and is governed by different factors than other types of analogy, specifically proportional analogy.

⁶For a discussion of the literature surrounding this question, see Fertig (2013: 71ff.)

⁷See Chapter 4 specifically for arguments in favor of this position.

1.4 PREVIOUS APPROACHES TO MORPHOLOGICAL LEARNING

Most studies of analogy over the last century and a half have not provided formal models for accounting for abstract morphological change. Rather, much of the focus has been given to laying out general principles from which historical linguists can substantiate their intuitions (cf. Paul 1886, Kuryłowicz 1966, *a.o.*). Even though these principles serve as the bedrock for our understanding of analogical change, it is worth examining more recent research that has been aimed at formalizing these intuitions. This section discusses some of the more prominent models, beginning with those which focus on productivity in Section 1.4.1, before discussing the MINIMAL GENERALIZATION LEARNER (Hayes & Albright, 1999) in Section 1.4.2. The MGL has been used to explain analogical change, simulating the process by which phonological and morphological generalizations can be learned. It has also been used in conjunction with notions of productivity, as I briefly discuss below.

1.4.1 APPROACHES INVOLVING PRODUCTIVITY

Productivity broadly refers to the capacity that speakers have to produce new and novel forms based on the rules and grammar that they have already learned. A simple look at derivational affixes in English allows us to make the generalization that we can make new words with the suffix *-ness* quite easily. However, we cannot as easily do the same with *-th*.⁸ Given the fact that *-ness* allows for the production of novel forms more easily than *-th*, it is clear that *-ness* is more productive than *-th*. Productivity, in this sense, is a property associated with specific morphemes or morphological patterns of expression.

There has been an ever growing emphasis on productivity with respect to how it interacts with analogical change. This should be of no surprise as analogical changes

⁸For the comparison between *-ness* and *-th* see Bauer 2001: 49.

easily lend themselves to the idea that older forms are simply overridden by productive constructions. Despite being framed as the result of specific proportional equations, proportional changes can often be stated as being the result of productive morphological patterns being applied to new items. Productivity in this sense refers not specifically to the comparison between A and B , but rather to the difference between A and A' as it is applied to new items. In this sense, the crucial factor which governs an analogy is the application of a pattern ($A \Rightarrow A'$), motivated in turn by its productivity. By modeling change this way, we can explain the analogy through a learning process which is not necessarily driven just by an arbitrary comparison between A and B specifically.⁹

To illustrate this using some of the examples from above, consider again the extension of the plural suffix $-s$ to *cow-s* (replacing *kine*). Instead of being based on an arbitrary comparison to a specific noun or set of nouns, the addition can simply be seen as *extension* of the productive plural ending. Doing so essentially restates four-part analogy as discussed above in Section 1.3.1 in terms of a rule whose application is motivated by the productivity of the plural marker.¹⁰

Similarly, the reinterpretation of *pease* as *pea+s* can be seen as part of the same productive process of the addition of $-s$ to nouns to produce plurals. The form *pease* is simply reanalyzed as yet another example of the already productive rule which adds $-s$ to nouns to form the plural. Analyzing the analogy this way has the added benefit of motivating both the changes affecting *cow-s* and *pea-s* under a single unified rule, which productively applies $-s$ to a stem identical to the singular. In this way, productivity

⁹Modeling the change in terms of productivity can thus help answer the question as to why we are comparing A and B in the first place.

¹⁰Stating the analogy in terms of a rule goes against Paul’s original formulation of proportional analogy, as his framework implied that morphology works on word-and-paradigm basis (Fertig, 2013). However, I see the distinction as a relatively minor one, as four-part analogies which assume complete word forms can generally be stated with functionally equivalent rules. Regardless of whether one assumes a word-and-paradigm approach or a rules-based approach, the abstraction required to extend the plural form with $-s$ on the basis of other nouns is essentially identical.

allows us to unify different types and instances of analogical changes as being motivated by the same pressure.¹¹

The diachronic formation of both *cows* and *peas* do not require a reference to specific nouns as models on which to base the analogical change. Instead, we can simply assume that speakers are applying a regular process to items that did not previously undergo that process. Thus, from a synchronic perspective, specific proportions serve more of a descriptive purpose rather than an explanatory one, as it does not provide an explicit motivation for the analogy. A productive rule/pattern, on the other hand, is extended on the basis that learners are biased against unproductive rules.¹²

The next few sections are devoted to discussing some of the recent proposals for modeling analogical change. I begin by describing two models for productivity, a statistical approach based on the work of Baayen (1989), and a learning model based on productivity known as the Tolerance Principle (Yang, 2016).

1.4.1.1 STATISTICAL EXPLANATIONS OF PRODUCTIVITY

An influential perspective to defining morphological productivity is taken by Baayen (1989, 1992, 1993), who uses a quantitative approach. Here productivity is defined in terms of the frequency of *hapax legomena*, a phrase which refers to words that only occur once in a delimited corpus. The basic assumption behind this work is that *hapax legomena* are representative of productive morphology being applied to novel forms.

¹¹The choice to extend the plural form *-s* as opposed to extending a different type of pluralization strategy must equally be reckoned with by word-and-paradigm approaches and rule-based approaches to proportional analogy. Productivity effectively is what is responsible for what comparison is being made. The change affecting A and A' is compared with B on the basis of how productive the change is. In this sense, productivity must be assumed by virtually any possible theory of proportional analogy, regardless of the nature of the morphological representations.

¹²This is not to say that there are never cases in which a four-part analogy can be used as an explanatory tool. It is certainly clear that in some cases specific lexical items influence others in terms of their morphological expression. However, whether it is the same mechanisms that govern these as those which produce the plural forms with an extended *-s* remains to be seen.

Specifically, the productivity of a given morphological process or category, notated as $\mathcal{P}$, is calculated as the ratio of n_1 to N , where n_1 is the number of types that occur only once, and N is the number of tokens of that morphological category. For example, if we are examining the different ways a language forms diminutives via suffixation, n_1 refers to the number of words which take a given suffix but only occur once in the corpus. This targets the hapax legomena which use the suffix in question. N in this case will refer to the sum of all instances in which the suffix is used as a diminutive. The ratio between the two measures how commonly the suffix is applied to new items, i.e. hapax legomena.

A secondary measurement is known as global productivity, notated as $\mathcal{P}^*$. It is calculated as the ratio of n_1 to the number of words that only appear once in the corpus, essentially expressing how meaningful the hapaxes of this morphological category are in comparison with how common other hapax legomena are in the corpus.

There are therefore two relevant measurements of productivity in this framework. The first, $\mathcal{P}$, measures how often a morphological category is used with words that occur only once. The second, $\mathcal{P}^*$, measures how often those hapaxes occur compared to other hapaxes.

It should be noted that, on their own, the quantitative measurements of productivity do not say much about analogical change. The quantitative variables can be used to determine whether one suffix is more productive than another, but they do not make any specific predictions regarding the learning of different competing affixes, which is presumably when analogy actually occurs. Thus, Baayen's approach is by itself insufficient in explaining analogical phenomena.

Sandell (2015) applies this quantitative framework to help evaluate some of the intuitions that specialists have regarding the productivity of different aorist types in

Ancient Greek and Sanskrit. In order to employ the quantitative benefits of measuring productivity, Sandell (2015) combines the use of Baayen’s modeling of productivity with the MINIMAL GENERALIZATION LEARNER (Hayes & Albright, 1999), expanded on in Section 1.4.2.

1.4.1.2 THE TOLERANCE PRINCIPLE

One of the more prominent proposals to come out of the recent literature concerning analogical change is that of Yang (2016). This among other works (Kodner, 2019, 2022) sets forth a type-based model by which rules can be learned depending on how many exceptions must be memorized. A general formula, known as the Tolerance Principle, has been offered to express this notion, seen below in (1).¹³

(1) **Tolerance Principle:** (Yang 2016: 64)

Let R be a rule applicable to N items, of which e are exceptions. R is productive if and only iff

$$e \leq \theta_N \text{ where } \theta_N := \frac{N}{\ln N}$$

The intuition behind this rule is that a rule is deemed to be productive if the number of cases where a rule should have applied but did not is lower than a certain threshold. The exact specifics of the rule in (1) are not important for this discussion. The crucial point is that a rule, R , is productive when there are few enough exceptions to the rule to allow for the learning of a rule (rather than rote memorization of particular alternating forms). How many exceptions are allowed depends on how many items could hypothetically undergo the rule. The more general the application of a rule is,

¹³As the intricacies of the mathematics behind the proposed equation are outside my area of expertise, I refer interested readers to the discussion in Yang (2016: 60-66).

the more exceptions are possible.

To illustrate this in general terms, consider again the English plural *-s*. The rule, namely add *-s* to the form in the singular to create the plural, is applicable to the vast majority of nouns. The exceptions, those nouns which for example take *-en* instead (*ox* → *oxen*), are few and far between. Thus, we can deem the English *-s* plural to be productive. The same cannot be said for the ‘add *-en*’ rule, whose forms must simply be memorized.

While the Tolerance Principle has generally been applied to different scenarios showing how children learn complex morphological distributions such as German plurals or gaps in Russian inflection rather than within the context of language change, it also provides a clear pathway for how and why analogical change might occur. Yang (2016) suggests that the Tolerance Principle is the reason for analogical leveling in *rew* (irregular past tense of *row*) → *rowed* (regular past tense) and analogical extension in *werede* (regular past tense of *wear*) → *wore* (irregular past tense based on *bear-bore*): ‘...both kinds of changes result from the application of rules to words that previously did not fall under their reign but nevertheless could have’ (Yang 2016: 157). In a similar vein, Kodner (2020) applies the Tolerance Principle to a hypothesized set of Proto-Germanic verbs in order to account for the presence of a lengthened **ē*-grade in the past of Class IV and V strong verbs, such as IV **beranq* ~ **bērun* and V **gebanq* ~ **gēbun*. In both proposals, the Tolerance Principle is the basis for determining which rules can be over-generalized. On the surface, such over-generalization leads to the introduction of well-established patterns to forms without such patterns, giving the impression of proportional analogies.

The Tolerance Principle can seemingly provide insights into proportional change, especially when the data allows for a quantitative analysis. However, the general principle

can also be applied in cases where there is not much data, as long as the productivity of the different patterns is clear.

1.4.2 OTHER LEARNING APPROACHES: MINIMAL GENERALIZATION

MINIMAL GENERALIZATION is an algorithm originally proposed by Hayes & Albright (1999) and extended by Albright & Hayes (2002), Albright (2002b,a), *etc.* that learns phonological and morphological rules based on progressive rule generation. Albright has used the algorithm to explicitly account for paradigmatic leveling, e.g. in Latin (Albright, 2002a) and Yiddish (Albright, 2010), and analogical extension in Spanish verbs involving diphthongization (Albright, 2009). Given the variety of analogical problems the MINIMAL GENERALIZATION algorithm has been applied to, it is worth summarizing the different aspects that constitute the algorithm.

There are two important parts to the algorithm, i) rule generation, and ii) rule reliability. The algorithm begins with pairs of forms, one of which is assumed to derive from the another. The forms are first listed as phonological strings, not yet broken down into their morphological constituents. The task of the algorithm is to generate rules that can derive one set from the other. Albright & Hayes (2002) illustrate this process with English verbs, pairing present tense forms with their past tense correlates, a subset of which are shown in (2):

- (2) English PRES~PAST (Albright & Hayes 2002: 2)
- ([mɪs]_{pres.}, [mɪst]_{past}) ‘miss(ed)’
- ([prɛs]_{pres.}, [prɛst]_{past}) ‘press(ed)’
- ([læf]_{pres.}, [læft]_{past}) ‘laugh(ed)’
- ([hʌg]_{pres.}, [hʌgd]_{past}) ‘hug(ged)’
- ([nɪd]_{pres.}, [nɪdəd]_{past}) ‘need(ed)’

After the comparison, the algorithm generates rules that account for the difference in these pairs, seen below:

(3) Rules generating English PAST TENSE (Albright & Hayes 2002: 2)

$$\emptyset \rightarrow t / \#mis_ \#$$

$$\emptyset \rightarrow t / \#pres_ \#$$

$$\emptyset \rightarrow t / \#laef_ \#$$

$$\emptyset \rightarrow d / \#h\Lambda g_ \#$$

$$\emptyset \rightarrow \text{əd} / \#nid_ \#$$

Once rules are generated which have the same structural change, the algorithm attempts to make generalizations. For example, (4) generalizes the fact that *-t* is added after stems ending in a sequence of a non-low, front, unrounded vowel and [-s]:

(4) Generalization between *miss* and *press* (Albright & Hayes 2002: 3)

$$\emptyset \rightarrow t / m \quad \text{ɪ} \quad s_ \#$$

$$+ \emptyset \rightarrow t / pr \quad \text{ɛ} \quad s_ \#$$

$$= \emptyset \rightarrow t / X \quad \left[\begin{array}{l} +\text{syllabic} \\ -\text{low} \\ -\text{back} \\ -\text{tense} \\ -\text{round} \end{array} \right] \quad s_ \#$$

The rule in (4) is highly specific to *miss* and *press*, leading to the name MINIMAL GENERALIZATION. This process is iterated many times, leading to a list of rules specified in terms of phonological contexts that account for the mapping between the pairs with the result that /t/ surfaces after voiceless obstruents.

The second important aspect of the algorithm involves the evaluation of the rules in terms of how they best account for the data. For each rule, a value of reliability

is assigned, determined by how often the rule derives the correct form compared to how many forms the rule could apply to. The reliability of each is adjusted depending on the sample size, also known as the confidence of the rule.¹⁴ Thus, the MINIMAL GENERALIZATION algorithm yields two parts, i) a list of rules, and ii) a value for each rule denoting how reliable that rule is based on its scope and success rate.

The reliability values allow rules to be compared. Thus, when presented with a novel input, the most reliable rule can be applied to produce a novel output. In this sense, the algorithm can be used to model analogy based on rules already learned by the grammar. As mentioned above, Albright has already extended the algorithm to various examples of analogical change, including in Latin (Albright, 2002a), Yiddish (Albright, 2010), and Spanish (Albright, 2009).

One question that arises is how the pairs which the rules generate are determined. In other words, what form in a paradigm serves as the base? Albright (2002a) proposes that the base in a paradigm can be determined by evaluating which form is *most informative* with respect to deriving the other forms in the paradigm. The most informative form is essentially the form whose rules have the highest reliability in deriving the other forms. To illustrate this, compare a hypothetical verb paradigm with six cells. Using the MGL, we can create six grammars, in which each form serves as the base from which the others are created. To determine the base, we can then compare these grammars to find the one which has the most reliable rules. When presented with novel forms, we can then use the rules determined from the most informative base to create new verb forms.

How does this algorithm relate to the discussion of proportional analogy specifically? As the MG algorithm determines a base by which other forms can be derived, it motivates

¹⁴This brief summary does a disservice to the full process. There are other small aspects that help improve the confidence involving phonological adjustments. A final grammar may also prune redundant rules. For more details, see Albright & Hayes (2002).

which cells in a paradigm a proportional analogy can target. Recall that a four-part analogy assumes the existence of a relationship between A and A' . What determines what A and A' can be? MINIMAL GENERALIZATION provides information about what A and B are, as those are the bases to which A' and B' are derived from. In this sense, the MG algorithm can inform on which analogies may be plausible.¹⁵

To summarize, the MINIMAL GENERALIZATION algorithm provides a means by which we can determine the morphological rules which can generate a paradigm. Furthermore, the framework proposes a method to determine which cell in a paradigm serves as the base: the most informative cell with respect to all the other cells.

1.5 MAIN CLAIMS

In terms of types of analogy, this dissertation focuses primarily on leveling and proportional change. Specifically, the main goal of the dissertation is to elucidate some of the factors that influence these two types of morphological change. Given that analogy, even when restricted to these two types, encompasses a large amount of data and many varieties of examples, I limit the scope of this dissertation to the *non-root intra-paradigmatic alternations* introduced in Section 1.2.

These alternations are not easily stated in terms of a four-part analogy. We will see more examples of these alternations in the next chapter; however, all are difficult to describe in these traditional terms. In regular examples of proportional analogy, a pattern abstracted from $A \rightarrow A'$ is extended to B' (e.g. $-s$ affecting both *stone-s* and *cow-s*). However, in the comparison between *andare* and *finire*, the difference between the exponents is arbitrary: $va(d) \sim finisc$ and $and \sim fin$. Nothing about $va(d)$ implies the addition of the segment *isc*.¹⁶ In this sense, there is no four-part proportion

¹⁵Analogies in which the paradigmatic base is not referenced are therefore presumably unmotivated.

¹⁶The same holds for any other formulation of the proportion.

that accounts for the relationship between the paradigms of the verbs *finire* and *andare*. Instead of an analogy expressed in terms of the exponents of the cells, the two paradigms are analogous in terms of the paradigmatic distribution.

Theories of productivity and learning algorithms have not yet been applied to the extension of such patterns of distribution. Chapter 2 will expand on how these frameworks interact with these patterns. First, however, it is necessary to detail the main claims made by the subsequent chapters.

1.5.1 CLAIMS

This section outlines the main claims made by this dissertation below. These claims roughly correspond to the different chapters, although many of the concepts are touched on throughout.

1. **Non-root Intra-Paradigmatic Alternations (Chapter 2):** The first goal of this dissertation is to describe the various types of *non-root intra-paradigmatic alternations* (see Section 1.2) and to make the case for further study of this phenomenon. While other types of morphological phenomena have received typological investigation, e.g. for suppletion see Veselinova (2006), Corbett (2007, *a.o.*), these alternations have been understudied as a unified phenomenon from a decompositional perspective.¹⁷

These alternations are of interest from both diachronic and synchronic perspectives. Diachronically, framing changes in terms of non-root alternations may augment the description of specific changes. However, the study of these alternations goes far beyond simply describing changes. Specifically, cases in which alterna-

¹⁷In some ways, the literature on stems does touch on these alternations, although rarely from an intra-paradigmatic perspective. Additionally, as I discuss in the next few chapters, there are theoretical and diachronic reasons to assume a decompositional approach to these alternations.

tions are created through analogy represent sophisticated examples of analogical change. As we will see, these changes often require a high level of abstraction on the part of a learner. The study of these alternations, therefore, can provide insight into what learners take into account when acquiring an alternation and probe the upper limits of abstract change. In other words, examining these alternations will make clear how sophisticated analogy of morphological phenomena can be.

From a synchronic perspective, these alternations are examples of seemingly highly idiosyncratic phenomena. Given that humans must be able to acquire such idiosyncratic morphological distributions, examples of such phenomena inform our understanding of what our morphological components need to (and do not need to) be able to encode.

2. Morphological alternations are governed by abstract generalizations based in the morphological component of grammar (Chapters 3 + 4):

As we will see, morphological alternations can generally be categorized by their paradigmatic distributions. Many examples fall neatly into clear morphosyntactic patterns. For example, one might find one alternant in the singular and another in the plural. However, there are also examples in which the distributions of the alternants are not so easily described in terms of a simple set of morphosyntactic features. In these cases, these alternations are often part of a larger morphological pattern in which different types of morphological phenomena occur in the same idiosyncratic distribution. This type of pattern has previously been described as *morphomic* in the sense of Maiden (2003, 2004, 2018, *etc.*), based on the original formulation by Aronoff (1994).

While the alternations which fall into neat distributions and those which do not

seem distinct on the surface, I will argue that both types make reference to underlying generalizations in the morphological component of internal grammar. As we will see in a case study about a subset of Latin irregular verbs, analogy makes direct reference to these generalizations, allowing for the proliferation of a synthetic-analytic alternation.

3. **Leveling (the spread of one alternant over another) in idiosyncratic distributions is governed by a set of linguistic biases (Chapter 4):** These biases include usage, frequency, and a desire to have morphological coherence.

As we will see in Chapter 4, alternations targeting disjunctive environments may be leveled over time. The process by which this leveling occurs is influenced by the aforementioned factors. In terms of usage-based biases, frequency directly impacts the acquisition of these irregular patterns. If something interrupts the frequency of specific cells, idiosyncratic patterns may be leveled.

Additionally, I argue that learners aim to internalize their morphological input in terms of generalizations. If frequency or use impacts this input by impoverishing certain paradigmatic cells, learners may level the alternants referencing those cells in favor of creating a more cohesive morphological generalization.

4. **A secondary goal is to show that morphemes are decomposable and despite surface irregularity are secondarily derived from more regular distributions.** Morphemes (see Section 2.4) are characterized by idiosyncratic distributions that must be listed in terms of disjunctive rules and environments in the morphological component of grammar. While it has been argued that examples of morphemes constitute evidence for morphological templates (cf. Maiden 2018), I argue that these ‘templates’ actually constitute a set of several disjunct rules,

whose sum is the morpheme, as evidenced by how morphomic distributions change over time.

In this sense, what separates morphomic distributions from distributions based on morphosyntactic features (like singular vs plural) is that morphomic distributions constitute the sum of several regular distributions.

While I adopt the framework of Distributed Morphology (Halle & Marantz, 1993) to formalize these structures, I hope to show that many of the key observations can be made with only general reference to the mechanics assumed by that framework.

In sum, this dissertation investigates two types of analogical change, proportional change (with a particular emphasis on morphomic analogies), and paradigmatic leveling, through the phenomenon of *non-root intra-paradigmatic alternations*. In focusing on these two phenomena in particular, this dissertation i) provides a novel empirical typology of several understudied types of analogical change, and ii) identifies several usage-based and structural biases in the form and application of analogical change.

1.6 ROADMAP

Chapter 2 begins by discussing *non-root intra-paradigmatic alternations* in more depth before providing a typology of the diachronic origins of different sub-types of these alternations. Among these alternations are highly irregular patterns, known as morphemes. Paradigms with these distributions provide the primary focal point of the subsequent chapters, as they represent the most challenging distributions to describe, both from a synchronic and diachronic perspective. Chapter 3 proceeds to discuss the role of morphomic structures in creating intra-paradigmatic alternations through the in-depth investigation of a case study concerning the synthetic-analytic alternations in

the paradigm of the Latin verb *nōlō* ‘I do not want’. The last portion of this chapter is dedicated to accounting for morphomic distributions from a synchronic perspective, comparing how two contemporary frameworks resolve the problematic distribution of athematicity in Latin irregular verbs. Chapter 4 investigates how analogical leveling affects a morphomic pattern in Romance, with special attention given to the alternations involving the suffix *-sc-*. Multiple developmental trajectories provide evidence that frequency, usage, and morphological coherence act as biases which can induce and condition leveling. The final chapter concludes, discussing the greater ramifications of the points made in Chapters Chapter 2-Chapter 4.

2

Non-Root Intra-Paradigmatic Alternations

This chapter examines *non-root intra-paradigmatic alternations* in depth and provides the backdrop for the subsequent discussions of morphomic alternations in Chapter 3 and Chapter 4. I begin by considering the object of inquiry in Section 2.1 and define what types of alternations *non-root intra-paradigmatic alternations* refer to. In doing so, I establish the position of these alternations as *bona fide* morphological phenomena as compared with other types of alternation, e.g. suppletion. Section 2.2 segues into a typology, enumerating the characteristics that define each type of alternation. We

can largely distinguish between phonological and morphological alternations, although both of these types can be subdivided in terms of various criteria. Section 2.3 provides examples of each type, further outlining their idiosyncrasies.

Afterwards, I introduce the notion of irregular distributions, known as morphemes. These alternations are not easily explained via either phonological or morphosyntactic mechanisms and pose challenges from both synchronic and diachronic perspectives. I argue that non-root intra-paradigmatic alternations can inform our understanding of how irregular patterns are encoded as well as how they develop over time. Section 2.4 expands on the diachronic aspects of these highly idiosyncratic distributions, while Section 2.5 discusses the relevance that these alternations have for synchronic theory.

2.1 DEFINING NON-ROOT INTRA-PARADIGMATIC ALTERNATIONS

This section expands on the explanation given in Section 1.2 regarding the nature of *non-root intra-paradigmatic alternations*. There are two relevant parts to this term, namely *non-root* and *intra-paradigmatic*. I will deal with each in turn.

2.1.1 NON-ROOT ALTERNATIONS

Non-root alternations refer to differences in morphological expression affecting a morpheme which is not the root. The most common type of non-root alternation involves affixes; an affix expressing a certain meaning alternates within a paradigm despite the grammatical function of the exponents remaining the same throughout the paradigm.

To illustrate this basic type, consider the following verbal paradigm from Romanian:

	SINGULAR	PLURAL
1ST	<i>lucr-éz</i>	<i>lucr-Ø-ám</i>
2ND	<i>lucr-éz-i</i>	<i>lucr-Ø-áți</i>
3ND	<i>lucr-eáz-ă</i>	<i>lucr-eáz-ă</i>

Table 2.1: PRES IND of *lucra* ‘to work’ in Romanian (Gönczöl 2020: 100)

In this paradigm, a suffix which generally surfaces as *-ez-* is responsible for marking the present tense; compare 1SG PRES IND *lucrész* with 1SG IMPF IND *lucram* (see Section 2.1.2). However, this suffix does not appear uniformly throughout the paradigm. In the 3SG/PL, the suffix appears as *-eaz-*, while the suffix is completely absent from the 1/2PL, shown as *-Ø* in Table 2.1. The suffix *-ez-* thus alternates between several allomorphs in a single paradigm. All of these allomorphs mark the present, and their form depends on the combination of person and number of the form in question. Crucially, the root of the verb, in this case *lucr-*, is unaffected by the alternation.

Such alternations are not limited to the verbal domain. Consider the derivational alternation in the Russian noun *medvežónok* ‘bear cub’ in Table 2.2:

	SINGULAR	PLURAL
NOMINATIVE	<i>medvež-ónok</i>	<i>medvež-át-a</i>
ACCUSATIVE	<i>medvež-ónk-a</i>	<i>medvež-át</i>
GENITIVE	<i>medvež-ónk-a</i>	<i>medvež-át</i>
DATIVE	<i>medvež-ónk-u</i>	<i>medvež-át-am</i>
INSTRUMENTAL	<i>medvež-ónk-om</i>	<i>medvež-át-ami</i>
LOCATIVE	<i>medvež-ónk-e</i>	<i>medvež-át-ax</i>

Table 2.2: *medvežónok* ‘bear cub’ in Russian

The suffix *-on(o)k/-at* creates derived nouns which denote the offspring of a given animal. In the singular, it appears as *-onok* in the nominative and as *-onk* elsewhere and may cause roots to undergo various phonological changes.¹ In the plural, the

¹These changes include changes in the final consonant of the root, movement of stress, and vowel deletion in the root (*jer* deletion). See Gouskova & Bobaljik (2022) for a more explicit account of the phonological behavior of this suffix.

exponent of the suffix has a radically different form: *-at*. Thus the suffix, while denoting the same semantic meaning throughout the paradigm, alternates in overt phonological shape dependent on number, as in *-on(o)k/-at*, and to a lesser extent case (*-onok* in the NOM SG and *-onk* elsewhere in the singular).² Once again, the root, *medved'* ‘bear’, is unaffected by the alternation, which only affects the suffix.

Morphological alternations which do not involve the root are not limited to affixal alternations. A second type of alternation that similarly affects non-root morphology occurs when periphrastic (or ANALYTIC) forms alternate with non-periphrastic (or SYNTHETIC) forms within a singular paradigm. For example, consider the present indicative of the Latin verb *nōlō* ‘I do not want’ (as compared to its forebear *volō* ‘I want’):

	NŌLLE ‘to not want’	VELLE ‘to want’
1SG	<i>nōlō</i>	<i>volō</i>
2SG	<i>nōn vīs</i>	<i>vīs</i>
3SG	<i>nōn vult</i>	<i>vult</i>
1PL	<i>nolumus</i>	<i>volumus</i>
2PL	<i>nōn vultis</i>	<i>vultis</i>
3PL	<i>nolunt</i>	<i>volunt</i>

Table 2.3: PRES IND of *nōlle* ‘to not want’ and *velle* ‘to want’ in Latin

The forms in Table 2.3 show an alternation between periphrastic forms (those with the free standing negative element *nōn*) in the 2SG, the 2PL, and 3SG and non-periphrastic forms in the 1SG, 1PL, and 3PL.

While the type of morphological expression is not strictly identical with the affixal type seen in Romanian (Table 2.1) and Russian (Table 2.2), it is comparable as the alternation specifically governs *non-root* morphology. For the examples of affixal suppletion, the fact that the root is not involved is clear. However, the periphrastic alternations similarly demonstrate this attribute. For *nōlō*, it is not the forms of the root which al-

²The latter alternation is phonologically motivated, while the alternation between *-onok* and *-at* is conditioned morphologically.

ternate.³ Rather, it is the appended material, the negation, which alternates depending on the person and number of the form.

As with affixal alternations, periphrastic alternations are also not limited to the verbal domain. For example, the Serbian Prizren-Timok dialect has examples of synthetic nominal forms alternating with analytic forms. Such an alternation is demonstrated below:

	SINGULAR	PLURAL
NOMINATIVE	<i>žena</i>	<i>žene</i>
ACCUSATIVE	<i>ženu</i>	<i>žene</i>
DATIVE	<i>ženī</i>	<i>na žene</i>

Table 2.4: *žena* ‘wife’ in Prizren-Timok, Svinitsa (Sobolev 2009: 723)

The paradigm in Table 2.4 shows a synthetic-analytic alternation based on case and number. In most of the forms, case is marked with an ending suffixed to the root *žen-*. In the dative plural, however, the accusative plural is used as a general oblique following the preposition *na* ‘on’. The replacement of synthetic nominal forms with analytic constructions containing either a preposition or postposition is, of course, not surprising. In Indo-European at least, the transition from morphological case to adpositional phrases is quite common (cf. Hewson & Bubenik 2006). What is peculiar in this case, however, is that the availability of a synthetic form in the dative is limited to the singular, creating a synthetic-analytic alternation in the nominal domain. It should be noted, once again, that the root remains unchanged; it is the morphology responsible for marking number and case which alternates in the paradigm in terms of periphrasis.

³See Chapter 3 for more details. Upon closer inspection, we may note that the 2SG form *is* suppletive; however, this is the case regardless of whether the form is negated or not (cf. 2SG *viś* ‘you want’), indicating that the suppletion of the 2SG is independent of the negation.

2.1.1.1 WHY NOT THE ROOT?

It is worth briefly touching on why there is such a focus in this discussion on alternations that do not involve the root. What is it about roots that warrants their omission?

The first reason is empirical. Alternations involving the root generally fall under the term *suppletion*, although some limit suppletion to alternations between phonologically unrelated exponents (e.g. Mel'čuk 1994, Bobaljik 2012).⁴ Suppletion as a phenomenon has been widely studied, from both synchronic and diachronic perspectives. While Section 2.1.4 expands on some of the research done in this area, we may note the existence of several typological studies with a focus on suppletion (Veselinova, 2006, Corbett, 2007, Bobaljik, 2012, *a.o.*). Likewise, from a diachronic perspective, the study of suppletion has a long history (Osthoff, 1899, Juge, 1999, 2019, Kim, 2019, *a.o.*). It cannot be said that alternations that do not affect the root enjoy the same amount of attention. In fact, even if we focus just on affixal alternations, there are very few studies that even consider framing their distributions in the way done here (cf. Section 2.1.4). Thus, giving specific attention to morphological phenomena which do not involve the root is worth doing simply because it has not been looked at before in depth.

Secondly, there are theoretical reasons to treat alternations involving roots and those which do not differently. Across different synchronic frameworks, it has been a long-held assumption that roots and affixes behave differently. For example, it has been argued from a morphological perspective that roots are unique in structural terms (cf. Halle & Marantz 1993, De Belder & Van Craenenbroeck 2015, Embick 2021). Even in morphological theories which do not specifically reference roots, the mechanisms which spell out stems also differ from those which spell out affixes. For example, in

⁴In this dissertation, I do not make a large distinction between suppletive pairs like *go~went* from pairs like *foot~feet*.

PARADIGM FUNCTION MORPHOLOGY (Stump 2001, 2015), a specific *Stem* function is responsible for the insertion of different stems, while affixes are added to stems via rule blocks. Phonological theories occasionally also make a theoretical distinction between affixes and roots based on morphological and phonological properties; see, for example, the ROOT-AFFIX FAITHFULNESS METACONSTRAINT (McCarthy & Prince, 1995).⁵

Given that there seem to be theoretical differences between roots and affixes that are borne out in how these two types of morphemes surface synchronically, it is worth considering how roots and affixes interface with analogical change. On the one hand, it may be possible that both affixes and roots are governed by the exact same diachronic principles, making a distinction between the two impractical from a historical perspective. On the other hand, it might also be the case that the two show distinct patterns of change and should therefore be treated differently. Either way, a dedicated treatment of non-root morphology from a diachronic perspective can be insightful in terms of setting up such a comparison.

2.1.2 LOCAL PARADIGMS AND INTRA-PARADIGMATIC ALTERNATIONS

Alternations that are *intra-paradigmatic* refer to those whose scope is restricted to a specific range. I have designated the term ‘local paradigm’ to describe the limited paradigmatic slices which this dissertation examines.

Descriptively, these paradigms denote not the full paradigm of a given lexeme, but rather the immediate paradigm that affects the alternation in question. The paradigm for the present indicative of the Romanian verb *lucr-* ‘work’ seen in Table 2.1 serves as an example of a local paradigm. The alternation of *-ez/-eaz/-Ø* is found within this singular paradigm, in which the forms only differ in terms of person and number.

⁵See also the discussion in Ussishkin & Wedel (2002), which argues that there are markedness and frequency differences between roots and affixes.

However, we may contrast the paradigmatic slice in Table 2.1 with the paradigm in Table 2.5, which shows an extended paradigm of the verb.⁶

	PRES IND	PRES SUBJ	IMPF	SMP PST	PLPF
1SG	<i>lucr-ez</i>	<i>lucr-ez</i>	<i>lucr-am</i>	<i>lucr-ai</i>	<i>lucr-asem</i>
2SG	<i>lucr-ez-i</i>	<i>lucr-ez-i</i>	<i>lucr-ai</i>	<i>lucr-ași</i>	<i>lucr-aseși</i>
3SG	<i>lucr-eaz-ă</i>	<i>lucr-ez-e</i>	<i>lucr-a</i>	<i>lucr-ă</i>	<i>lucr-ase</i>
1PL	<i>lucr-Ø-ăm</i>	<i>lucr-Ø-ăm</i>	<i>lucr-am</i>	<i>lucr-arăm</i>	<i>lucr-aserăm</i>
2PL	<i>lucr-Ø-ați</i>	<i>lucr-Ø-ați</i>	<i>lucr-ați</i>	<i>lucr-arăți</i>	<i>lucr-aserăți</i>
3PL	<i>lucr-eaz-ă</i>	<i>lucr-ez-e</i>	<i>lucr-au</i>	<i>lucr-ară</i>	<i>lucr-aseră</i>

Table 2.5: Extended paradigm of *lucra* ‘to work’ in Romanian

By examining the extended paradigm in Table 2.5, we can make the following generalization. The alternation between *-ez-* and *-Ø-* only appears in the present tense, while the alternation between *-ez-* and *-eaz-* is limited to the present indicative; it is not found in the other tenses, which either demonstrate no affixal alternations or demonstrate their own affixal alternations that are independent in form and distribution from that of the present. By only focusing on the local paradigm, we can hone in on the cells where the alternation seems to be meaningful.⁷

If we did not restrict the scope of the alternation to the immediate local paradigm, one might draw a comparison between the present indicative/subjunctive stems and the stems in the other tenses. In this case, we would be comparing the presence and absence of *-ez-* throughout the entire paradigm. However, such a comparison conflates the nature of the alternation within a local paradigm with that in an extended paradigm. The suffix *-ez-* is absent from the imperfect, not because of a morphological idiosyncrasy that we

⁶I have left out the imperative in Table 2.5 due to concerns of space. Its presence and absence is immaterial to the general point made in this section.

⁷If we apply this logic to the distribution in Romance, the present indicative and subjunctive should manifest themselves in two local paradigms. While this is indeed the case, I have treated them together here as they both show the same alternation. As we will see in Chapter 4, the present indicative and present subjunctive can behave independently from one another.

need to account for specifically, but rather because *-ez-* is a present tense marker and therefore not expected to appear in the imperfect tense. The same cannot be said of the 1/2PL of the present indicative and subjunctive, where we would expect *-ez-* to appear. This dissertation focuses on alternations that occur in local paradigms rather than those that occur more broadly, as the former represent cases of morphological irregularity and thus are of a greater interest from both diachronic and synchronic perspectives.

Imposing limits on what part of a paradigm we examine necessarily brings up the question of how we determine what exactly local refers to in a specific context. I offer two practical criteria that allow us to define the scope of a local paradigm. Firstly, the meaning of the morphological exponents that engage in the alternation suggests which part of the paradigm is relevant. For example, if an affix spells out the present tense, it makes sense to narrow the scope of the paradigm to the present tense, as the absence of the affix in a past tense would be expected.⁸

Secondly, alternations that occur within local paradigms as discussed in this dissertation are conditioned by features that are structurally external to the morpheme.⁹ Essentially, morphemes that occur farther away from the affix condition the alternating allomorphs. On the other hand, morphemes that occur inside the structure, such as the root, do not condition the allomorphy. In this sense, the features responsible for the alternations are limited. For verbs this refers mainly to person and number, while for nouns this refers to number and case. For the example in Table 2.1, the appearance of *-ez-* is dependent on two aspects: i) the tense, and ii) the person and number of the form in question. The first is expected as this is what the suffix refers to, while the second conditions the alternation within the paradigm. In other words, it is the 1PL

⁸The suffix *-ez-* above is an example of a present tense marker.

⁹I do not intend to make a theoretical claim here. Rather, from a descriptive point, we can use this directionality as a criteria to limit the scope of the alternations we are examining.

and 2PL that cause *-ez-* to be absent in the present, while forms like the 1SG PRES IND cause *-ez-* to be present.

2.1.3 INTERIM SUMMARY

In sum, this dissertation focuses on intra-paradigmatic alternations that involve specifically non-root morphology. In terms of non-root morphology, there are two types, i) those involving affixes (surfacing as either prefixes or suffixes), and ii) periphrasis, where the alternation may consist of free standing elements and non-free standing elements. The crucial unifying factor is that these alternations do not involve the exponence of the root.

The other defining characteristic of these alternations is that they occur within a paradigm. The scope of such alternations is limited to a *local paradigm*, defined in this section in terms of what the affix or periphrasis expresses as well as what features condition their exponence. As we will see, these alternations often represent examples of morphological irregularity and thus are of interest in terms of their diachronic origin and their synchronic representation.

2.1.4 MOTIVATING AFFIXAL SUPPLETION

Previous research has largely ignored non-root intra-paradigmatic alternations in terms of a unified approach. This is not surprising, given that many of the alternations that this term could apply to are largely categorized under various disparate morphological phenomena. For example, affixal alternations are often described in terms of straightforward *allomorphy*, with little to no emphasis on the intra-paradigmatic dimension of the alternation. Periphrasis is handled as a completely separate phenomenon, often with a more syntactic perspective.

Although patterns of suppletion in terms of roots or stems have been well-studied within the literature (Mel'čuk, 2006, Corbett, 2007, Bobaljik, 2012, Smith et al., 2019, *a.o.*), alternations that affect non-root morphology are not usually framed in these terms. The term *affixal suppletion* has been used by Mel'čuk (1971, 1976, 1994, 2006) to describe allomorphy which specifically targets affixes. This term has received some criticism (e.g. Corbett 2007), given the inclusion of inflectional allomorphy which reflects declension and conjugation class.¹⁰ However, I will use the term *affixal suppletion* to refer to intra-paradigmatic alternations in which the exponency of affixes is conditioned not by the root, but by features that are expressed in a more outward position than the affix in question.

I suggest that affixal alternations often resemble alternations involving root suppletion. One need look no further than the Russian diminutive suffix *-onok* (1), as outlined above in Table 2.2. Recall that in the singular, the suffix has the form *-onok/-onk*, The plural suffix, however, has the shape *-at*, creating an alternation SG *-onok/-onk* ~ PL *-at*.

- (1) Diminutive suffix *-onok* in Russian (Mel'čuk 2006: 426)

l'vj-onok 'lion cub' → *l'vj-at* 'lion cubs'

medvež-onok 'bear cub' → *medvež-at* 'bear cubs'

We may thus justify the use of *suppletion* in terms of the alternation in (1) on theoretical grounds. The difference between the forms with *-onok* and those with *-at* is based on a feature (number) that is peripheral to the affix. In Russian, these are expressed on

¹⁰In response to Mel'čuk's use of the term suppletion in these contexts, Corbett (2007: 15) argues that this use conflates two different phenomena, namely root suppletion and inflectional allomorphy, stating 'if affixal suppletion is allowed for, then suppletion ceases to be a special boundary phenomenon requiring specific analysis (and a special term) and rather takes in a substantial proportion of inflectional morphology'. The criticism of the term therefore relates specifically to its use to describe affixal allomorphy whose conditioning environment is dependent on the root.

the inflectional endings that attach to the root + affix complex. It is this which is what conditions the affix suppletion, not some characteristic of the root.¹¹ The fact that more peripheral features condition affixal suppletion also directly parallels cases of root suppletion which are likewise conditioned by peripheral features. A clear example, also from Russian, is *čelovek*- ‘person’ which has a suppletive plural based on the stem *ljud*- in the plural.

	SINGULAR	PLURAL
NOMINATIVE	<i>čelovék</i>	<i>ljúdi</i>
ACCUSATIVE	<i>čelovéka</i>	<i>ljuděj</i>
GENITIVE	<i>čelovéka</i>	<i>ljuděj</i>
DATIVE	<i>čelovéku</i>	<i>ljúdjam</i>
INSTRUMENTAL	<i>čelovékom</i>	<i>ljud’mí</i>
LOCATIVE	<i>čelovéke</i>	<i>ljúdjax</i>

Table 2.6: *čelovek* ‘person’ in Russian

The crucial observation here is that the allomorphy in both the alternation in (1) and the forms in Table 2.6 is conditioned by something external to the root, i.e. number. In this sense, allomorphy of the type seen in (1) directly parallel intra-paradigmatic suppletive alternations such as the present tense of English ‘to be’ (*am/is/are*), where the person and number conditions the exponency of the root.

Exactly how the similarities between (1) and Table 2.6 are formalized is a matter of choice of framework. However, it is clear that there is some merit to assuming that similar mechanics are at play in both of these cases, suggesting that suppletion as a theoretical notion may also play a role in describing affixal alternations.

¹¹This contrasts with other types of affixal allomorphy which may mark declension or conjugation class. Here, internal morphology such as the root or stem conditions the allomorphy.

2.2 DIACHRONIC TYPOLOGY OF ALTERNATIONS

As we will see, not all alternations are equally relevant for the discussion of analogy. Some are derived due to phonological changes, either diachronically or synchronically. Others are inherently morphological in nature, or at the very least do not have a simple phonological conditioning environment. This section lays out the different types of non-root intra-paradigmatic alternations based on their diachronic origin.

It is necessary to identify the diachronic sources of these intra-paradigmatic for several reasons. Firstly, we must be able to distinguish between alternations which are phonological in nature from those which are morphological. It is the latter which is of most interest from a historical perspective that focuses on constraining analogical change. Secondly, having an idea of where these alternations come from cross-linguistically can help us assess the original distribution of the alternants in a given alternation. This is important as an investigation of leveling (cf. Chapter 4) requires knowing the original distribution. Lastly, in those cases in which an external element has been introduced to a paradigm, a few examples of which we will see shortly, we want to know what the source of those extra-paradigmatic elements can be. In short, establishing a diachronic typology allows us to have a better understanding of the foundational patterns of change that affect intra-paradigmatic alternations, whether it be how alternations come to be or how they intersect with analogy.

The section proceeds as follows. I will provide several examples of each type; before doing so, however, I briefly discuss some preliminaries regarding the general categorization of the different types.

2.2.1 GENERAL CATEGORIZATION

The different types of non-root intra-paradigmatic alternations can largely be separated into two main groups based on their diachronic history. The first group consists of examples of whose alternation is based in terms of phonology. This type can be further divided in terms of whether the alternation is phonologically active within a given language’s synchronic grammar. As we will see, phonologically conditioned alternations often become grammaticalized, depending on whether the synchronic rules which govern the alternation remain active or not.

The second group consists of cases where the alternation is based solely on morphological change. This group can be further divided into several sub-types. At this point, we can introduce the terms *INVASIVE* and *NON-INVASIVE*. The former refers to cases where some or all members of an alternation have been introduced to a paradigm externally from different diachronic sources. The latter, on the other hand, consists of those cases where the alternants are traceable to a singular historical origin. Thus, we can designate two types of alternations whose origin is morphological, depending on whether the source of the alternation is external or internal. Crucially, this classification is ultimately dependent on the origin of the relevant morphemes and says nothing specific about the synchronic status of these alternations.

The comparison between different suppletive alternants in terms of their diachronic origin is, of course, not new.¹² The term *INVASION* is similar to the notion of *incursion* as proposed in Juge (1999, 2013, 2019). Here, *incursion* is used to describe cases of non-affixal suppletion where the forms of one word intrude on another. We have already seen a famous example of *incursion*, namely the conflation of the reflexes of *vādere* ‘to

¹²The comparison of different intra-paradigmatic affixal and periphrastic alternants from a diachronic perspective is new.

rush’ and *ambulāre* ‘to walk’ in Romance, as seen in Table 2.7. Here, the forms of *vādere* have *invaded* several of the cells in the present tense.¹³

	ITALIAN	CATALAN	FRENCH	OLD SPANISH
1SG	<i>vàdo</i>	<i>vaig</i>	<i>vais</i>	<i>voy</i>
2SG	<i>vài</i>	<i>vas</i>	<i>vas</i>	<i>vas</i>
3SG	<i>và</i>	<i>va</i>	<i>va</i>	<i>va</i>
1PL	<i>andiàmo</i>	<i>anem</i>	<i>allons</i>	<i>imos</i>
2PL	<i>andàte</i>	<i>aneu</i>	<i>allez</i>	<i>ides</i>
3PL	<i>vànno</i>	<i>van</i>	<i>vont</i>	<i>van</i>

Table 2.7: Root Suppletion Alternations for ‘to go’ in Romance (Maiden 2004: 232-233)

The focus on suppletive alternations with etymologically distinct origins has also played a key role in previous typologies of non-affixal suppletion (cf. Osthoff 1899). Rudes (1980) defines true suppletion when roots are etymologically distinct (as opposed to ‘pseudo-suppletion’). Also similar is the distinction between ‘proper suppletion’ and ‘phonological suppletion’ put forth by Börjars & Vincent (2011).¹⁴

While the notion expressed by the term INVASIVE is largely compatible with previous typological schemas, there is at least one reason why I choose this term over others. Apart from the term incursion, other terms describing suppletive alternations do not describe the nature of their etymological origins, besides the fact that they are distinct. The term INVASIVE directly refers to the fact that some cells in a local paradigm have been targeted by introduction of external morphology. Other schemas simply do not commit to this sort of directionality. While the term incursion as defined by Juge (2019) is similar, its scope does not map on fully to what we see in terms of the alternations discussed in this chapter.

¹³We can infer that *vādere* has invaded the paradigm given that the reflex of *ambulāre* is found throughout the non-present tenses. See Chapter 4 for a more in-depth discussion.

¹⁴See also the discussion of suppletion in Kim (2019).

2.2.2 SUMMARY OF THE DIFFERENT TYPES

The different types of non-root intra-paradigmatic alternations can be schematized as in the tree in Figure 2.1. The following categories are based on the diachronic origin of the alternation.

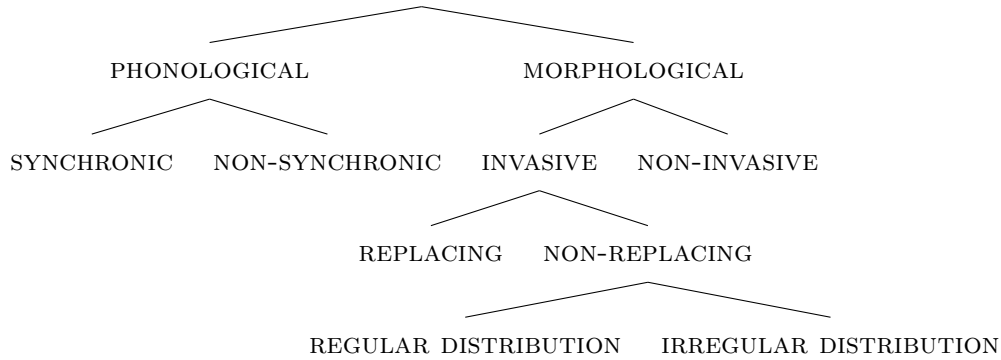


Figure 2.1: Types of non-root alternations

The next few sections go through the different types, providing examples of each. The examples are chosen on the basis of the diachronic clarity. Unless stated explicitly otherwise, the diachronic developments discussed below are uncontroversial.¹⁵

2.3 TYPES OF NON-ROOT ALTERNATIONS

2.3.1 PHONOLOGICAL

Phonological alternations can be subdivided into two types, depending on whether the alternation is still synchronically active in phonological terms.

¹⁵In other words, I will explicitly state wherever a novel claim is made with regards to the specific diachronic developments.

2.3.1.1 SYNCHRONICALLY ACTIVE ALTERNATIONS

Affixal alternations can be shown to be synchronically active if there is a phonological conditioning environment which can be justified on independent grounds. This section provides several examples of such alternations. However, beyond this discussion, I will largely put aside such alternations as they are not relevant for an investigation analogical change.¹⁶

LENGTH ALTERNATIONS IN LATIN VERBS: To express the imperfect tense in Latin, the suffix *-bā-* is added to a stem consisting of a root and a theme vowel, before the personal endings corresponding to person and number are added. To illustrate its distribution, consider the imperfect of the verb *amāre* ‘to love’ below in Table 2.8:

	SINGULAR	PLURAL
1ST	<i>amābam</i>	<i>amābāmus</i>
2ND	<i>amābās</i>	<i>amābātis</i>
3RD	<i>amābat</i>	<i>amābant</i>

Table 2.8: IMPF ACT IND of *amāre* ‘to love’ in Latin

The long vowel *-ā-* alternates with a short counterpart throughout the paradigm. Compare the long vowel in 2SG *amābās* with the short vowel in 3SG *amābat*. The differences in length are synchronically dependent on the consonant which follows the vowel. The exact phonological environment can best be described in historical terms. Ultimately, the synchronic restrictions of vowel length are due to the culmination of several diachronic phonological sound changes: i) vowels were shortened before final *-m* (1SG *amābam*), ii) vowels were shortened before a sequence of a sonorant consonant followed by another consonant, also known as Osthoff’s law (3PL *amābant*), and iii) vowels were shortened before final *t* and *r* in polysyllabic words (3SG *amābat*).¹⁷

¹⁶That said, it is important to know which alternations can be set aside.

¹⁷For a description of these changes, see Weiss 2020: 136-140.

That the suffix is synchronically understood to be *-bā-* with a long vowel underlyingly is supported on the grounds that the suffix engages in different alternations that also involve length, conditioned by the same phonological environments. The passive imperfect also utilizes the suffix *-bā-*, although in this paradigm, the alternation is slightly different, as demonstrated in Table 2.9.

	SINGULAR	PLURAL
1ST	<i>amābar</i>	<i>amābāmur</i>
2ND	<i>amābāris</i>	<i>amābāminī</i>
3RD	<i>amābātur</i>	<i>amābāntur</i>

Table 2.9: IMPF PASS IND of *amāre* ‘to love’ in Latin

Here, the vowel is short before the final *-r* in the 1SG as well as the 3PL which, like the 3PL of the active, is subject to shortening due to Osthoff’s law. If the alternations in Table 2.8 and Table 2.9 were morphological, a Latin speaker would have to memorize the specific cells in which the shortened variant appeared; if the alternation is interpreted as a side effect of regular phonological processes, such specific memorization is not needed, as the environments can be stated generally. An analysis in which the shortening is interpreted phonologically is simply more economical than the alternative.

Additionally, other affixes in Latin also show similar shortening effects, including long theme vowels, e.g. the first conjugation marker *-ā-* (3SG *amat* contrasting with 2SG *amās*), and the present subjunctive marker *-ē-* (1SG *amem* vs 2SG *amēs*). It is clear from these examples that it is their phonological environment which unites the instances in which shortening occurs; the alternative, namely that each instance is listed morphologically, is implausible.

CONSONANT QUALITY IN THE IMPERFECT IN OLD CHURCH SLAVIC: Another example of a phonologically-induced alternation that is synchronically active, rather than morphologically conditioned, is the alternation between *x~š* in the imperfect marker

-*ěax-* in Old Church Slavonic.

	SINGULAR	DUAL	PLURAL
1ST	<i>zov-ěax-ъ</i>	<i>zov-ěax-ově</i>	<i>zov-ěax-omъ</i>
2ND	<i>zov-ěaš-e</i>	<i>zov-ěaš-eta</i>	<i>zov-ěaš-ete</i>
3RD	<i>zov-ěaš-e</i>	<i>zov-ěaš-ete</i>	<i>zov-ěax-ъ</i>

Table 2.10: IMPF IND of *zъvati* ‘to call’ in Old Church Slavonic

As with the Latin imperfect, the conditioning environment is clearly phonological. Here, the consonant *-x-* alternates with *-š-* depending on the vowel that follows. Before a front vowel, *-x-* is palatalized to *-š-*, while *-x-* is preserved before non-front vowels.

This process of palatalization is part of a larger system of phonological alternations in which the quality of consonants is affected by the vowels which follow. The changes that affect *-x-* above are directly paralleled by alternations involving other velar consonants ($k \rightarrow \check{c}$, $g \rightarrow \check{ž}$, $x \rightarrow \check{s}$).

For direct evidence that this process was still synchronically accessible in Old Church Slavonic in terms of $x \sim \check{s}$, note the following examples involving *-x-*, e.g. NOM SG *dux-ъ* vs VOC SG *duš-e* ‘spirit’. Similarly, the marker for the aorist *-x-* is similarly also governed by phonological restrictions. It demonstrates the same alternation before front and non-front vowels: 1SG *nesoxъ* versus 3PL *nesoše* ‘carry’.

VOWEL QUALITY IN THE $\bar{A}$ -PRETERITE IN TOCHARIAN B: A final example of a phonologically conditioned alternation is exhibited by the Class I Preterite in Tocharian B, where the suffix *-ā-* [a] alternates with *-a* [ə] in the third singular.

	SINGULAR	PLURAL
1ST	<i>sruk-ā-wa</i>	<i>sruk-ā-m(o)</i>
2ND	<i>sruk-ā-sta*</i>	<i>sruk-ā-s(o)</i>
3RD	<i>sruk-a</i>	<i>sruk-ā-re</i>

Table 2.11: PRET IND of *sruk-* ‘die’ in Tocharian B

This alternation is conditioned by the synchronic assignment of stress affecting di-

syllabic words. The third singular, as a disyllabic form, regularly has the accent on the root syllable, while the other forms are accented on the syllable containing *-ā-*.

The synchronic conditioning of these changes is confirmed by the fact that when the third singular is followed by a clitic which adds an additional syllable, the stress shifts rightward to the syllable with *-a-* [ə], subsequently surfacing as *-ā-* [a], e.g. 3SG *lyukā-me* ‘lit them up’.

SUMMARY: These examples demonstrate alternations which are easily stateable in phonological terms (lengthening, palatalization, vowel quality) and whose conditioning environments can likewise be stated in phonological terms (when preceding certain consonants or vowels, when stressed, etc.). The phonological changes are paralleled by other independent examples in the respective languages, demonstrating that the nature of the alternation is phonological rather than morphological. While it is useful to be familiar with these examples, especially as they interact with other types of alternations, I will not focus on them further in any detail.

2.3.1.2 NON-SYNCHRONICALLY ACTIVE ALTERNATIONS

Moving on to non-synchronically active alternations, this type is characterized by differences conditioned by regular phonological change. The difference between alternations that are synchronically active and those that are not is that either the phonological change is no longer accessible or the forms involved have simply been morphologized as the exponents producing the alternation. Often the difference between the alternants can still be stated in phonological terms, although this is ultimately dependent on the phonological changes involved.

LENGTH ALTERNATIONS DUE TO SZEMERÉNYI'S LAW IN ANCIENT GREEK: Szemerényi's law refers to a well-established sound change that affected PIE in which a sequence of $*-VRs$, where R refers to a resonant, was subject to the deletion of the $-s$ accompanied by compensatory lengthening of the vowel preceding the resonant.¹⁸ This change can be illustrated by the outcomes of the NOM SG of $*ph_2tér$ 'father' which ultimately went back to $**ph_2ters$. The outcome of the long vowel created by the compensatory lengthening can still be seen in the daughter languages: Gk. $\pi α τ ή ρ$ *patēr*, Skt. *pitā* 'father'.

The NOM SG of stems ending with liquids and nasals was a perfect environment for this sound change, given that the general ending for the NOM SG was $*-s$. Upon the deletion of the $-s$ and lengthening of the vowel, these nominative singulars were now characterized by a long vowel: e.g. OLat. *uxōr* 'wife' (Pl. *Asin.* 927) and PIE $*s_{iu}(H)méns$ 'membrane' > $*s_{iu}(H)mén$ > Gk. $\acute{\upsilon} μ ή ν$ *humén*. Accordingly, this sound change also lengthened the nominative singular of words with the agent noun suffix $*-tor$, best reflected by Gk. $-τωρ$ *tōr* < $*tor-s$. This lengthened variant alternates with the form $-τορ$ *-tor* in the other cells in the paradigm. Table 2.12 demonstrates how the agent suffix, appended to the root $dō-$ 'give' < $*deh_3-$, alternates in length.

	SINGULAR	PLURAL
NOMINATIVE	$\delta \acute{\omega} - τ ω ρ$ <i>dō-tōr</i>	$\delta \acute{\omega} - τ ο ρ - ε ς$ <i>dō-tor-es</i>
ACCUSATIVE	$\delta \acute{\omega} - τ ο ρ - α$ <i>dō-tor-a</i>	$\delta \acute{\omega} - τ ο ρ - α ς$ <i>dō-tor-as</i>
GENITIVE	$\delta \acute{\omega} - τ ο ρ - ο ς$ <i>dō-tor-os</i>	$\delta \omega - τ ό ρ - ω ν$ <i>dō-tór-ōn</i>
DATIVE	$\delta \acute{\omega} - τ ο ρ - ι$ <i>dō-tor-i</i>	$\delta \acute{\omega} - τ ο ρ - σ ι$ <i>dō-tor-si</i>
VOCATIVE	$\delta \ddot{\omega} - τ ο ρ$ <i>dō-tor</i>	$\delta \acute{\omega} - τ ο ρ - ε ς$ <i>dō-tor-es</i>

Table 2.12: $\delta \acute{\omega} τ ω ρ$ *dōtōr* 'giver' in Ancient Greek

The difference in length is a direct result of the sound change. The forms that show a short $-o-$ are expected, given Szemerényi's Law would not have applied in these

¹⁸For a description of the sound change, see Fortson (2011).

circumstances.

While the sound change may have been part of a synchronic process in PIE, there is little evidence to suggest that that is still the case in these examples. There is also no other synchronic phonological motivation for the lengthened outcome in Ancient Greek. For example, one might imagine some phonotactic restriction on a sequence of $-VR\#$ with a short vowel; in this sense, the lengthening would be a synchronic phonological repair. However, the VOC SG form $\delta\tilde{\omega}\tau\omicron\rho$ *dōtor* confirms that no such lengthening process is at play. Thus, speakers had to simply encode the fact that the nominative singular form contains a long vowel in the morphology.

ROMANCE VERBS WITH SUFFIXAL ALTERNATIONS: One class of French verbs is characterized by the vowel $-i-$ which alternates with the exponent $-is-$ throughout the paradigm. Here, we will focus on the alternation in the present indicative. Consider the paradigm of the verb *finir* ‘to finish’:

	SINGULAR	PLURAL
1ST	<i>fin-is</i> /fin-i/	<i>fin-iss-ons</i> /fin-is-õ/
2ND	<i>fin-is</i> /fin-i/	<i>fin-iss-ez</i> /fin-is-e/
3RD	<i>fin-it</i> /fin-i/	<i>fin-iss-ent</i> /fin-is/

Table 2.13: PRES IND of *finir* ‘to finish’ in French

The paradigm in Table 2.13 demonstrates an alternation between forms with $-is(s)-$ and those without. Modern French orthography suggests that the comparison is between the 3SG and the other cells; however, this merely obscures the true alternation between $/-i-/$ in the singular and $/-is-/$ in the plural. It is unlikely that an underlying $-s$ is simply deleted in the singular, given the presence of $-s$ in the 3PL.¹⁹ Thus, underlyingly there is an alternation between $-i-$ and $-is-$ that must be expressed morphologically, rather than phonologically.

¹⁹There do exist accounts which propose one exponent which underlies both the forms in the singular and those in the plural, e.g. Fenne (2020).

While the alternation between /-i-/ and /-is-/ is not phonological in nature in Modern French, the origin of the alternation was clearly conditioned by phonological changes. Beginning with the forms in Proto-Romance, seen below in Table 2.14, we can see the Latin verbal suffix *-sc-* was found with no alternation throughout the paradigm.²⁰

	SINGULAR	PLURAL
1ST	* <i>finīscō</i>	* <i>finiscimus</i>
2ND	* <i>finīscis</i>	* <i>finiscitis</i>
3RD	* <i>finīscit</i>	* <i>finīscunt</i>

Table 2.14: PRES IND of **finiscō* ‘I finish’ in Proto-Romance (Lanly 1977: 329)

A crucial observation from Table 2.14 is that the final endings of the plural forms contained substantially more phonological material than those of the singular. Both the 1PL and 2PL had an extra syllable compared with the singular forms, while the 3PL had a final consonant cluster. This extra material in effect protected the *-sc-* from various phonological changes involving syncope and several rounds of apocope involving the deletion of final consonants. This development is confirmed by intermediate forms of the 3SG, e.g. Old French *finist* and Middle French *fini(s)t*, where the form has been largely modified from the original **finīscit*, but preserves the *-sc-*. The singular of the Modern French forms have lost the final consonant, producing [fini] across the board.

This section will not go into the exact phonological developments in detail, but the general pathway towards an alternating /-i-/~/-is-/ is clear. Several rounds of phonologically-regular deletion have reduced the *-is-* in the singular (along with their final desinences) to *-i-*, creating a morphological alternation along the lines of number.

A similar case in Romance occurs in Romanian. Here, the inherited distribution of verbs with the suffix *-ez-* < Lat. *-idi-* was also in part affected by phonological developments.²¹ Recall the distribution of *-ez-/eaz-* in the Romanian verb *lucr-* ‘work

²⁰See Chapter 4 for more information on the Latin suffix and its reflexes in Romance. Appendix B.3.1 also discusses the French scenario in more detail.

²¹See Chapter 4.

from above, restated below:

	SINGULAR	PLURAL
1ST	<i>lucr-éz</i>	<i>lucr-Ø-ăm</i>
2ND	<i>lucr-éz-i</i>	<i>lucr-Ø-ăți</i>
3ND	<i>lucr-eáz-ă</i>	<i>lucr-eáz-ă</i>

Table 2.15: PRES IND of *lucra* ‘to work’ in Romanian (Gönczöl 2020: 100)

Putting aside the alternation between the suffix *-ez-* and *-Ø-*, the difference between *-ez-* in the first and second singular and *-eaz-* in the third person forms is the result of a regular phonological process which has been morphologized.

Vowel breaking due to stress is a common process in Romance languages and is pervasive in Romanian. Historically, vowels like *-e-* were diphthongized when stressed and when they were not followed by a high vowel (Maiden et al. 2020: 15). From a synchronic perspective, stress still causes some phonologically conditioned alternations in Romanian: *beat* /beát/ ‘drunk’ ~ *bețú* /betsív/ ‘drinker’ (Chitoran 2002: 209). The diphthongal outcome in *beat* is the result of the vowel breaking when stressed. However, diphthongal outcomes have been morphologized throughout the language, confirmed by the distribution of *-ez-/eaz-*. For example, the 1SG shows a monophthongal outcome, *lucréz*, despite being stressed. From a diachronic perspective, the monophthongal outcome is expected due to the fact that at one point, the following syllable contained a high vowel *-u* < Lat. *-ō*. Similar outcomes can be seen in other verb types in Romanian (Chitoran, 2002, Maiden et al., 2020).

SYNTHETIC-ANALYTIC ALTERNATION IN OLD CHURCH SLAVONIC: An example of a synthetic-analytic alternation which was created through phonological motivations is demonstrated by the negated form of *byti* ‘to be’ in Old Church Slavonic. This verb shows a contraction in the present tense which alternates on the basis of the initial sequences of the verb.

	SINGULAR	DUAL	PLURAL
1ST	<i>něsmb</i>	<i>něsvě</i>	<i>něsmъ</i>
2ND	<i>něsi</i>	<i>něsta</i>	<i>něste</i>
3RD	<i>něstъ</i>	<i>něste</i>	<i>ne sptъ</i>

Table 2.16: PRES IND of *ne byti* ‘to not be’ in Old Church Slavonic

These forms are the result of a contraction between the negative element *ne* and the present tense of *byti*, whose forms begin with *e-* [je-] throughout the present tense, e.g. *esmb*, *esi*, *estb*, etc. apart from the third plural which begins with an *s-*: *sptъ*, where no contraction takes place. Thus, the distribution of uncontracted forms (the 3PL) and contracted forms (elsewhere) followed an expected phonological distribution.²²

THEME VOWEL ALTERNATIONS IN ANCIENT GREEK (O-CLASS): The thematic declension characterized by *-o-* in Ancient Greek demonstrates a complicated set of alternations, involving both phonological and analogical changes. We will focus on the distribution of this noun class in the Attic dialect:

	SINGULAR	PLURAL
NOMINATIVE	$\acute{\alpha}\nu\theta\rho\omega\pi\text{-}\omicron\varsigma$ <i>ánt^hrōp-os</i>	$\acute{\alpha}\nu\theta\rho\omega\pi\text{-}\omicron\iota$ <i>ánt^hrōp-oi</i>
ACCUSATIVE	$\acute{\alpha}\nu\theta\rho\omega\pi\text{-}\omicron\nu$ <i>ánt^hrop-on</i>	$\acute{\alpha}\nu\theta\rho\omega\pi\text{-}\omicron\upsilon\varsigma$ <i>ánt^hrop-ous</i> [-u:s]
GENITIVE	$\acute{\alpha}\nu\theta\rho\acute{\omega}\pi\text{-}\omicron\upsilon$ <i>ánt^hrōp-ou</i> [-u:]	$\acute{\alpha}\nu\theta\rho\acute{\omega}\pi\text{-}\omega\nu$ <i>ánt^hrōp-ōn</i>
DATIVE	$\acute{\alpha}\nu\theta\rho\acute{\omega}\pi\text{-}\omega$ <i>ánt^hrōp-ōi</i>	$\acute{\alpha}\nu\theta\rho\acute{\omega}\pi\text{-}\omicron\iota\varsigma$ <i>ánt^hrōp-ois</i>

Table 2.17: $\acute{\alpha}\nu\theta\rho\omega\pi\omicron\varsigma$ *ánt^hrōpos* ‘man’ in Attic

The distribution of the thematic vowel *-o-* in these forms suggests that the thematic vowel and the final desinences could be decomposed into independent affixes; however, in practice, a full decomposition of these affixes proves difficult. Particularly problematic are the GEN SG, ACC PL, and GEN PL, as the boundary between thematic vowel and desinence has been obscured by historical developments. Regardless, I offer a potential

²²Depending on the phonological analysis that one takes, this example could be analyzed as a synchronically active alternation. See Section 2.3.1.3.

decomposition that we might assume below:²³

	SINGULAR	PLURAL
NOMINATIVE	- <i>o-s</i>	- <i>o-i</i>
ACCUSATIVE	- <i>o-n</i>	- <i>u:-s</i>
GENITIVE	- <i>o-o</i> → - <i>ou</i> [u:]	- <i>o-ōn</i> or - <i>∅-ōn</i>
DATIVE	- <i>ō-i</i>	- <i>o-ís</i>

Table 2.18: Potential decomposition of the thematic declension in Attic

If we decompose the thematic vowel and desinences as in Table 2.18, we are left with an alternation between -*o-*, -*u:-*, and -*ō-*.²⁴ Decomposing these complexes in this way is controversial. It is certainly possible that synchronically many of these endings were frozen at the point of attestation and thus were no longer accessible to speakers. However, I believe that the decomposition posited in Table 2.18, or something similar, was accessible to speakers at some point in time. The evidence for this assumption is that the decomposition of the sequence of thematic vowel + desinence is implied by outcomes in α -stems which hinge on analogies with the *o*-stems. For example, the NOM PL -*oi* < *-*oḡi* exerted analogical influence on the ending of the α -stem NOM PL -*ai* -*ai* which was remodeled from original *-*ā-es* < *-*eh₂-es*. This analogy implies that both the theme vowels -*o-* and -*α-* were treated abstractly by those speakers who facilitated the analogy. It would otherwise be difficult to motivate the intrusion of the ‘a-quality’ of the vowel.²⁵

Although ultimately the alternation between -*o-* and -*ō-* is also an example of a

²³For the prehistory of these endings, see Sihler (1995), Weiss (2020).

²⁴Some of these decompositions may be altered depending on the assumptions made. For example, the thematic vowel in the GEN PL could easily be posited as -*∅-*, given that -*ων* -*ōn* is also used for the other noun classes. Additionally, a clear possibility is that for many of the endings the sequences of thematic vowel + desinence were interpreted together as a portmanteau morpheme, with no specific reference to its constituents. This was apparently already the case for some of these endings in PIE. See discussion below. Despite these issues, the general observation that the thematic vowel is not uniform throughout the forms that were decomposable remains. It is worth mentioning that the segmentation of the endings is not only motivated by the identification of the thematic vowel. This segmentation also serves to bring this noun class in line with other noun classes which also utilize the desinences to the right of each complex, cf. α -stems. See also the discussion in Section 2.3.2.1.

²⁵Section 2.3.2.1 expands on this point.

phonological alternation that is no longer synchronically conditioned,²⁶ in this section I focus specifically on the ending of the ACC PL in Attic, -ους *-ous*, as it demonstrates an interesting case in which an alternant was introduced due to a diachronic sound change.

The underlying ending of the accusative plural of thematic nouns in Proto-Greek can clearly be reconstructed to **-ons*, which is directly attested in several dialects, e.g. Central Cretan τὸνς ἐλευθέρωνς *tons eleuthéronς* ‘the free men’ (*IC* 4 72 *Col.* 7.7-8).²⁷ What sets the Attic form apart from these dialects is that it has undergone the second compensatory lengthening, by which **-n* was lost in the sequence **-Vns* with lengthening of the preceding vowel. Thus, **-ons* became **-ōs* in those dialects in which final *-ns* was subject to the secondary compensatory lengthening.²⁸ The Attic form -ους *-ous* [-u:s] is the result of a raising of *-ō-* to [u:]. The alternation between [o] and [u:] is thus the culmination of these two sound changes: i) the second compensatory lengthening, and ii) the raising of *-ō-*.

This alternation is unique, as other dialects of Ancient Greek that did not undergo the second compensatory lengthening do not show alternations in the accusative plural, instead preserving the original *-ns*. Even those that participated in sound changes that affected the sequence **-ns#* maintain *-o-* in the accusative plural. For example, Elean participates in a process of ‘compensatory diphthongization’, producing accusative plurals such as ἐπιστίοιρ *ep^histioir*.²⁹ Even here, the thematic vowel is apparent on the surface.

The introduction of alternations to the thematic declension is by no means unique

²⁶The long vowel in the DAT SG was due to a contraction of **-o-ei* (cf. Sihler (1995: 258), Weiss (2020: 239)), although this alternation was of much older date than the vowel change in the ACC SG

²⁷Central Cretan also shows *n*-less accusative plurals like the article in the phrase τὸς καδεστάνς *tos kadestans* (*IC* 4 72 *Col.* 7.43-44) which are the result of a sandhi rule in which the cluster *-ns#C* was simplified. This sandhi change was also operational at the Proto-Greek stage.

²⁸/o/ refers to a slightly raised [o].

²⁹Elean has independently undergone rhoticism of the final *-s*, cf. Kühlert (2021).

to Ancient Greek. The nominative plural ending of the second declension in Latin, $-\bar{i}$ (*lup̄i* ‘wolves’), notably lacks any surface sequence which resembles $-o-$, unlike the NOM SG *lupus* and ACC SG *lupum*, where the $-u-$ preceding the final consonant is the regular outcome of the thematic vowel. Behind the NOM PL $-\bar{i}$ form underlies $*-o\bar{i}$, which has become $-\bar{i}$ through a series of regular sound changes (cf. Weiss 2020: 240). This example, alongside forms like the GEN PL in Ancient Greek, suggest a trend in which the thematic vowel is lost due to regular sound changes accompanied by varying degrees of reanalysis of the sequence thematic vowel + desinence into a portmanteau morpheme.

SUMMARY: The examples outlined above demonstrate how phonological alternations can become morphologized. This process affects a whole host of different types of morphological categories, including verbs and nouns, and can also be shown to produce synthetic-analytic outcomes.

2.3.1.3 MORPHOLOGIZATION

A fundamental issue arises when comparing synchronically accessible alternations with those that have been morphologized. As is clear from the examples above, many of examples of morphologized alternations resemble plausible synchronic alternations. We can observe that the morphologized alternations tend to still be describable in phonological terms (length, diphthongization, deletion, vowel raising, etc.). For example, in both Section 2.3.1.1 and Section 2.3.1.2, there are alternations that involve vowel length (Lat. $\bar{a}/a$ and Gk. $o/\bar{o}$). This resemblance is not surprising, given that the diachronic process of sound change was often conditioned by synchronic phonology.

However, we must ask ourselves how we can ever be sure if an alternation has been morphologized. This question has become ever more pertinent, given the great number of abstract phonological apparatuses that have been proposed to explain what

traditionally have been considered to be morphological phenomena (e.g. noun class in Italian as in Lampitelli & Ulfsgjorninn (2023)). Is it possible to explain all of these alternations as synchronically accessible if we assume an abstract expanded phonological toolkit?

The recognition of this problem is not new. Notably, Kiparsky (1968) considers this very issue, arguing against completely abstract representations of phonological phenomena in favor of phonetically motivated underlying representations. Discussions of morphologization specifically also acknowledge this issue. For example, in the context of consonant mutation in Celtic, Janda (2017: 405-406) brings up the fact that the accounts that try to explain phenomena that are traditionally considered to be morphological solely via phonological mechanisms rely on stipulated assumptions.

This section does not aim to provide a solution to this problem nor does it give a method by which we can determine whether an alternation is synchronically accessible or not. I merely bring up this point to draw attention to the possibility that some of the examples described as having been morphologized may be explainable in phonological terms, albeit with the caveat that certain additional assumptions may be required. Ideally, such assumptions are supported independently. I assume, however, that there do exist alternations that are truly morphological in nature, as discussed in the next few sections.

2.3.2 MORPHOLOGICAL ALTERNATIONS

The following sections discuss the different types of morphological alternations. These are characterized by morphological change, rather than phonological change. These are of particular interest to us, as analogy plays a role in both creating the alternations as well as changing their distribution via leveling.

2.3.2.1 NON-INVASIVE (INTERNAL CHANGES)

Non-invasive alternations involve the development of alternants based on morphemes which are local to the paradigm in question. Essentially, no segment is introduced to the local paradigm from a (paradigm-)external construction. Crucially, this does not entail that there is no comparison between the local paradigm and an external paradigm. As we will see, analogy can create and extend alternations on the basis of patterns that exist elsewhere in the language without introducing an intrusive element.

ATTIC $\bar{\alpha}$ -DECLENSION ALTERNATIONS: The first declension in Attic Greek encompasses nouns that are characterized by a ‘theme vowel’ $\bar{\alpha}$ $-a$: which follows the root and precedes the final ending which marks number and case, mirroring the structure of the *o-class* discussed in Section 2.3.1.2. There are several different types, each with their own idiosyncrasies produced by a set of phonological and analogical changes. As we will see, these noun classes represent an insightful example of internal change. In what follows, I provide a novel framing of what are relatively uncontroversial changes to the plural endings.³⁰

Consider the following types of first declension noun in Ancient Greek below in Table 2.19:³¹

³⁰The developments affecting these endings are usually treated individually. Here, I frame their overall development together.

³¹The sequences involving V-V sequences are diphthongal on the surface, despite the morphological decomposition below. Thus, the dative plural ending *-ais* is a true diphthong [-a $\bar{\alpha}$ is].

	TYPE 1 χώρᾱ ‘land’	TYPE 2 νίκη ‘victory’	TYPE 3 μοῖρα ‘fate’	TYPE 4 γλῶττα ‘tongue’
NOM SG	χώρᾱ <i>k^hó:r-a:</i>	νίκη <i>ní:k-ε:</i>	μοῖρα <i>moîr-a</i>	γλῶττα <i>glô:tt-a</i>
ACC SG	χώρᾱν <i>k^hó:r-a:-n</i>	νικῆν <i>ní:k-ε:-n</i>	μοῖραν <i>moîr-a-n</i>	γλῶτταν <i>glô:tt-a-n</i>
GEN SG	χώρᾱς <i>k^hó:r-a:-s</i>	νικῆς <i>ní:k-ε:-s</i>	μοῖρᾱς <i>moîr-a:-s</i>	γλῶττης <i>glô:tt-ε:-s</i>
DAT SG	χώρᾱ <i>k^hó:r-a:-i</i>	νικῇ <i>ní:k-ε:-i</i>	μοῖρᾱ <i>moîr-a:-i</i>	γλῶττῃ <i>glô:tt-ε:-i</i>
NOM PL	χώραι <i>k^hô:r-a-i</i>	νῖχαι <i>nî:k-a-i</i>	μοῖραι <i>moîr-a-i</i>	γλῶτται <i>glô:tt-a-i</i>
ACC PL	χώρᾱς <i>k^hó:r-a:-s</i>	νικᾱς <i>ní:k-a:-s</i>	μοῖρᾱς <i>moîr-a:-s</i>	γλῶττᾱς <i>glô:tt-a:-s</i>
GEN PL	χωρῶν <i>k^ho:r-ô:n</i>	νιχῶν <i>ní:k-ô:n</i>	μοιρῶν <i>moir-ô:n</i>	γλωττῶν <i>glo:tt-ô:n</i>
DAT PL	χώραις <i>k^hó:r-a-is</i>	νίχαις <i>ní:k-a-is</i>	μοίραις <i>moîr-a-is</i>	γλῶτταις <i>glô:tt-a-is</i>

Table 2.19: $\bar{a}$ -declension showing alternations between [a], [a:], [ε:], and [∅] in Attic

The four types in Table 2.19, represented by column, each have a slightly different type of alternation. The different alternants are color coded depending on both vowel quality and length: [a:], [a], [ε:].³² I will deal with each type individually; however, before doing so, we can observe that these types are only distinguished in the singular. The plural for all types involve the same distribution of alternants, a short [a] in the NOM PL and DAT PL and a long [a:] in the ACC PL.³³

In the singular, TYPE 1 and 2 are uniformly characterized by a long vowel. TYPE 3 and 4 are split between the strong cases (NOM SG and ACC SG), where the theme vowel is a short [-a], and the weak cases (GEN SG and DAT SG) where there is a long vowel. In sum, all of the types have the same alternations in the plural, while TYPE 3 and 4 have a further alternation in the singular between long and short vowels.

We can largely generalize these alternations into three separate synchronic alternations, summarized as follows:

³²The GEN PL is marked with [∅].

³³For the purposes of this discussion, I put aside the genitive plural. The Attic form does not have a vowel before the long [o:] of the ending, although this is an Attic development. Other dialects preserve a vowel in this position; Homer has [a:], e.g. θεᾱών ‘goddesses’, also cf. Ionic Περσέων ‘Persians’. The Attic form is the result of contraction. While the GEN PL does seem peculiar from a synchronic perspective, it should be noted that the GEN PL of the *o*-class can also be analyzed with a -∅- before the ending -*o:n*, bringing both noun classes in line with one another.

- (2) Generalizations about the synchronic alternations
- a. Broadly, short [a] in the plural forms (apart from the accusative plural) alternates with a long vowel ([a:] or [ɛ:]) in the singular.
 - b. Long [a:] in the accusative plural alternates with short [a] in the nominative and dative plurals
 - c. Short [a] in the strong cases alternates with a long vowel ([a:] or [ɛ:]) in the weak cases in the singular of TYPE 3 and TYPE 4

Before discussing these alternations in further detail, it is important to note the original sources of the main distinctions between the different types. TYPE 1 and TYPE 2 are both descendants from the PIE nominal type characterized by the suffix **-eh₂-*.³⁴ This suffix was found throughout the paradigm. In other words, there were no alternations in the original distributions; all the alternations that are seen in Table 2.19 are due to inner-Greek developments. In Greek, this suffix is preserved as *-ᾱ* in most dialects. However, in Attic-Ionic dialects, **ᾱ* was fronted to **ē*. In Attic, the final result of this fronting is *ε:*. However, when the vowel followed *ε*, *i*, or *r*, **ē* reverted to **ᾱ*, which remained so in Attic. Thus, the **ᾱ* was preserved following *ε*, *i*, and *r*, while otherwise becoming *ε:*. This sound change is the source of the difference between TYPE 1 (preservation of **ᾱ*) and TYPE 2 (forms in which **-ᾱ* became *ε:*).³⁵

TYPES 3 and 4 share an entirely different origin from TYPE 1 and 2. Instead of the suffix **-eh₂-*, these types are characterized by the ablauting suffix **-ih₂-*/**-ieh₂-*, also known as the *devī* suffix.³⁶ The difference between the strong cases and the weak cases

³⁴This suffix is responsible for marking the first declension in Ancient Greek and Latin. This class mostly consists of feminines. The suffix itself is likely related to the ending of the neuter plural **-h₂* (Weiss 2020: 245).

³⁵For more information on these changes as well as the developments affecting the specific endings, see Sihler (1995).

³⁶See Fortson (2011).

in these types is attributable to the difference in ablaut: NOM SG $*-ih_2 > *-i\bar{a}$ compared to GEN SG $*-i\bar{e}h_2-es > *-i\bar{a}-es > *-i\bar{a}s$. The long [a:] and [ɛ:] found in the GEN SG and DAT SG are thus etymologically distinct from the vowels found in TYPES 1 and 2. In this sense, at least some of the similarities between the conjugations are coincidental, based on phonological convergence.³⁷ The difference between TYPES 3 and 4 is precisely the same as the difference between TYPES 1 and 2, in that TYPES 2 and 4 involves fronting of [a:] to [ɛ:].

Given the synchronic similarity between these declensions, it is tempting to try to explain some of the differences between the classes through synchronic phonology, as this would reduce the number of classes from four to two at the minimum.³⁸ However, it is difficult to motivate such a reduction, as it would be difficult to say what the underlying vowel would be. While we might be able to motivate a shift from [ɛ:] to [a:] after [ɛ, i, r], the forms in the accusative plural counter such an assumption. If the underlying vowel were [ɛ:], we would expect $*\nu\acute{\iota}\kappa\eta\varsigma$ [ní:kɛ:s] for the ACC PL. Any other underlying vowel would be difficult to motivate. While it may be possible to explain the difference via an account that assumes a significant amount of abstraction (see Section 2.3.1.3), I assume for the remaining discussion that there is a meaningful morphological difference between the different classes.

Returning to the intra-paradigmatic alternations, the differences in length that we see between the plural and singular cannot be accounted for by phonological processes. There is no synchronic rule in Ancient Greek that would motivate the shortening of [a:] or [ɛ:] to [a] before [i], as is the case in the nominative and dative plurals. In fact,

³⁷It should be noted that various changes affecting the sequence of consonant plus $*-i\bar{e}$ have obscured the presence of the $*-i\bar{e}$ in TYPES 3 and 4. Additionally, the $-a$ in the accusative singular of these types was analogically introduced from the NOM SG.

³⁸That is, one class that contains nouns in TYPE 1 and 2, and one class that contains nouns in TYPES 3 and 4.

the DAT SG and NOM PL of TYPES 1 and 3 demonstrate a minimal pair, in which the forms are only distinguished in terms of the length of the vowel: $\chi\tilde{\omega}\rho\alpha\iota$ $k^h\hat{o}:r\text{-}\mathbf{a}\text{-}i$ versus $\chi\acute{\omega}\rho\alpha$ $k^h\acute{o}:r\text{-}\mathbf{a}\text{-}i$.³⁹ Likewise, there is no lengthening process that would explain the long vowels in the singular. The idiosyncratic distributions of length and quality simply point to a distribution that is conditioned morphologically.

For the rest of this section, I will explain the source of the generalized alternations listed in (2). The idiosyncratic distributions of length in the paradigms seen on the surface in Attic obscure a relatively straight-forward history, which combines phonological and morphological developments to produce the alternations. As we will see, analogical change has affected these four types, particularly in the nominative plural, where a direct comparison with the *o*-class has allowed for the spread of the short [a] alternant.

Beginning with the dative plural, the shortened vowel in *-a-is* is the consequence of Osthoff's law,⁴⁰ in which a long vowel is shortened in the sequence $-VRC$.⁴¹ Thus, the short [a] seen in this form is due to regular sound change: $^*\bar{a}\text{-}is > \text{-}a\text{-}is$. Despite having a long vowel in the accusative plural, the ending *-a:s* actually also points to a regularly conditioned short vowel as in **-ans* with a short vowel, which was originally similarly conditioned by Osthoff's law: $^*\bar{a}ns > ^*ans$. I assume here that the $^*n\text{-}$ in the accusative plural was restored in the $\bar{a}$ -stems from an inherited $^*\bar{a}s < ^*eh_2ns$. The long vowel seen in Attic is the result of the second compensatory lengthening.⁴² This lengthening occurred after the change from $^*a:$ to $^*\epsilon:$, explaining the presence of *-a:* in TYPE 2. From a diachronic perspective, the ending *-a:s* in the accusative plural to

³⁹The differences in the accent are conditioned in this pair by the quality of the vowel. Crucially, the quality of the vowel is independent from the accent.

⁴⁰See the discussion on Latin verbs in Section 2.3.1.1.

⁴¹That is a sequence of vowel + resonant + consonant.

⁴²See the discussion of the *o*-class in Ancient Greek in Section 2.3.1.2.

conceals a set of regular phonological developments.

Both the accusative plural and dative plural point to a stage in which there was a short [a] in these forms alternating with a long [a:] in the rest of the paradigm. Thus, the fact that there is a short alternate in the plural is ultimately due to regular phonological change.

What about the nominative plural? Unlike the accusative and dative plural, there is no phonological pathway that accounts for the nominative plural. An analogical solution involving the comparison with the *o*-stems seems likely, as has been generally assumed in the literature (e.g. Sihler 1995). However, it is not exactly clear what the proportion would be; perhaps assuming the ACC PL as the base would be best: $*-o-ns : *-o-i :: *-a-ns : X$, $X = *-a-i$.

One way to motivate the change in the nominative plural is to assume that both the presence of a short [a:] in the ACC PL and DAT PL may have further facilitated the comparison. In other words, learners may have generalized a pattern in which the plural (apart from the GEN PL) is characterized by a short vowel.

Consider the following schematic development in Table 2.20 and Table 2.21:⁴³

⁴³Several notes have to be made regarding the schema below:

1. The stages below are rough representations of development. The changes between the different stages did not necessarily take place concurrently.
2. STAGE 1 roughly corresponds with a pre-Greek distribution, inheriting the regular PIE ending in the nominative plural. However, the *-n-* in the accusative plural of the *ā*-stems has already been restored (see above). Osthoff's Law has not yet taken place.
3. STAGE 3 corresponds to the distribution seen in Attic. Other dialects have may deviate slightly from this (e.g. the ACC PL).
4. The NOM PL during STAGE 2 has undergone an analogical replacement in both *o*- and *ā*-stems based on the pronominal endings.

	STAGE 1		STAGE 2		STAGE 3
NOM PL	*- <i>o-s</i>	>	*- <i>o-s</i> → *- <i>o-i</i>	>	- <i>oi</i>
ACC PL	*- <i>o-ns</i>	>	*- <i>o-ns</i>	>	- <i>ūs</i>
GEN PL	*- <i>ōn</i>	>	*- <i>ōn</i>	>	- <i>ōn</i>
DAT PL	*- <i>ō-is</i>	>	*- <i>o-is</i>	>	- <i>ois</i>

Table 2.20: Development of *o*-stem plurals

	STAGE 1		STAGE 2		STAGE 3
NOM PL	*- <i>ā-es</i>	>	*- <i>ā-es</i> → *- <i>a-i</i>	>	- <i>ai</i>
ACC PL	*- <i>ā-ns</i>	>	*- <i>a-ns</i>	>	- <i>ās</i>
GEN PL	*- <i>ā-sōm</i>	>	*- <i>ā-(h)ōn</i>	>	-(<i>ā</i>)- <i>ōn</i>
DAT PL	*- <i>ā-is</i>	>	*- <i>a-is</i>	>	- <i>aīs</i>

Table 2.21: Development of the *ā*-stem plurals

STAGE 1 represents the endings of the *ā*- and *o*-stems in pre-Greek. At some point, Osthoff's law shortened the *-ā-* theme vowel in the ACC PL and DAT PL, as well as the *ō* in the DAT PL of the *o*-stems, producing short vowel outcomes in these forms.

This change facilitated the comparison between the *ā*-stems and the *o*-stems in terms of length. Both declensions would have been characterized by a short vowel in the accusative and dative plurals, allowing for a direct analogy between the forms in the NOM PL. In other words, the *-o-* would have contrasted with *-ā-* throughout the singular and plural.

During STAGE 2, the pronominal ending *-i* was introduced to both the *o*- and *ā*-stems. In the *o*-stems, the replacement of *-s* and *-i* proceeded without a change in length. In the *ā*-stems, the vowel was short, although a long vowel would have been expected on the basis of the long vowel *-ā-* throughout most of the paradigm. I suggest that learners would have had a plausible reason to compare a short [a] with a short [o] as a marker of the plural, given the fact that elsewhere the plural of both declensions have a short vowel (besides the GEN PL). In this sense, the shortening in the *ā*-stem NOM PL was introduced on the basis of both the short vowel in the *o*-stem as well as the short [a] elsewhere in the *ā*-stem plural.

Finally, STAGE 3 represents the Attic forms shown in Table 2.19 after the second compensatory lengthening introduced a long vowel to the accusative plural in both declensions.⁴⁴

The alternations in TYPES 3 and 4 in the singular are due to an inherited difference in strong versus weak stems expressed by ablaut. In this sense, these forms simply inherit a morphological distinction that just so happened to be reflected in Greek by means of length (and in the case of TYPE 4, quality as well). The plural forms would have resembled the regular $\bar{a}$ -stems. Accordingly, the nominative plural of TYPES 3 and 4 also underwent the introduction of the pronominal ending, along with a shortening of the preceding [a:]. In this sense, TYPES 3 and 4 would have served to reinforce the changes seen in the regular $\bar{a}$ -stems.

To briefly reiterate the general point of this section, the short vowel in the NOM PL in the $\bar{a}$ -class in Ancient Greek arose due to analogical comparison with the *o*-class. This change can be framed as a rule whereby the vowel is short in the plural (apart from the GEN PL). This rule was extended to the various types of nouns with $\bar{a}$. What separates this analogical change is the no element has been introduced to the paradigm externally.⁴⁵ Rather, the *-a-* quality of the vowel is simply a retention of the general shape of the exponent present throughout the paradigm. In this sense, the development here is strictly internal and non-invasive.

This section has provided a novel framing of the changes affecting length in the $\bar{a}$ -stem plurals, in which various regular phonological developments have produced a situation where the plural of $\bar{a}$ -stems could be compared with that of *o*-stems.

⁴⁴Additionally, **o:* was raised in the *o*-stems. See the discussion about *o*-stems in Section 2.3.1.2.

⁴⁵I exclude here the addition of the pronominal ending *-i*, which is orthogonal to the alternation between *-ā-*/*-a-*.

THE SUFFIX *-IDI-* IN ROMANCE: Another example of an alternation produced by internal change is the paradigm involving the reflex of the Latin suffix *-idi-* in several Romance languages. We have already seen an example of this, namely the Romanian verb *lucr-* ‘to work’. Recall the suffix *-ez/-eaz* < Lat. *-idi-* is absent in the first and second plural. I will not go into much detail here, as I discuss this example and related cases in Chapter 4. Essentially, the deletion of the suffix in the 1PL and 2PL is due to an analogical comparison with other present classes, in a pattern known as the N-PATTERN. Crucially, there was no invasion of an external element.⁴⁶

2.3.2.2 REPLACEMENT

Contrasting with the non-invasive alternations discussed above, there are also several types of invasive alternations that occur when a cell within paradigm *A* is replaced by a cell from paradigm *B*, introducing an alternation involving non-root morphology. In such cases, the key attribute is that there is a complete replacement of the exponents and as such accompanied by any attributes unique to paradigm *B*.

LATIN FUTURE An example of an affixal alternation which demonstrates this change can be found in the third and fourth conjugations in Latin.⁴⁷ Consider the future active paradigm of the third conjugation verb *ferō* ‘I bear’ in Latin:

	SINGULAR	PLURAL
1ST	<i>fer-a-m</i>	<i>fer-ē-mus</i>
2ND	<i>fer-ē-s</i>	<i>fer-ē-tis</i>
3RD	<i>fer-e-t</i>	<i>fer-e-nt</i>

Table 2.22: FUT of *ferō* ‘I bear’ in Latin

The future in these forms is characterized by the suffix *-ē-* followed by the personal

⁴⁶Besides, of course, the original borrowing of the suffix *-idi-* from Ancient Greek *-ιζω*. However, the suffix was borrowed into the entire paradigm.

⁴⁷See Appendix B.1.2 for an overview of the Latin conjugation system.

endings, appearing throughout the paradigm apart from the 1SG where an *-a-* is found.⁴⁸ This alternation affects the third (see above) and fourth conjugations (1SG *audiam* ‘I will hear’) in both the FUT ACT and FUT PASS (1SG *ferar* ‘I will be brought’ and *audiar* ‘I will be heard’).⁴⁹ It is clear from the evidence in Old Latin that an original 1SG *-e(m)* consistent with the rest of the paradigm was replaced by *-am* (cf. Weiss 2020: 442).

So where is the *-a-* in the future from? The source is relatively clear. The 1SG of present subjunctive has replaced the 1SG of the future, such that both the future and present subjunctive are identical in form. For the whole paradigm of the present subjunctive, see below:

	SINGULAR	PLURAL
1ST	<i>fer-a-m</i>	<i>fer-ā-mus</i>
2ND	<i>fer-ā-s</i>	<i>fer-ā-tis</i>
3RD	<i>fer-a-t</i>	<i>fer-a-nt</i>

Table 2.23: PRES SUBJ of *ferō* ‘to bear’ in Latin

The replacement of future forms by the subjunctive forms is easy to motivate; the pathway from SUBJ → FUT is well-attested. After all, the Latin future was itself originally derived from an inherited subjunctive. The future marker *-bi-* seen in the first and second conjugations is a root aorist subjunctive of **b^huh_x-* ‘to be’, while the third and fourth declensions continue the thematic present subjunctives (Weiss 2020: 442). Modern colloquial Romanian uses a periphrastic construction based on the subjunctive to express the future (Maiden et al. 2020: 372), as is the case in other Balkan languages. What is unique about this specific case in Latin is why only the first person singular was replaced rather than the entire paradigm.

I suggest that the introduction of the subjunctive to the first singular future has a

⁴⁸The long vowel *-ē-* alternates with a short *-e-* due to regular phonological changes. See Section 2.3.1.1 for a discussion on vowel length alternations in Latin.

⁴⁹The first and second conjugations utilize the suffix *-bi-* in their respective future forms: *amābis* ‘you will love’, *monēbis* ‘you will advise’. Notably, the ending of the 1SG is different in these conjugations than the ending seen in the third and fourth; we find *-ō* rather than *-m*: *amābō/monēbō*.

usage based explanation. It is likely that different person and number configurations may be associated with different uses, especially in the subjunctive, given how many different functions may be expressed by this mood. What is required for the replacement of the 1SG in the future is some affinity for the 1SG subjunctive to express precisely such future meaning.

There is some evidence for such an assumption. Cross-linguistically, there are examples of first person modals which have a future-like functionality. For example, as pointed out by [Leech et al. 2009](#): 80-81, *shall* can be used as a future in the first person (*Then shall we start?*) in British English. This usage, however, is dispreferred for the other persons. *Shall* is acceptable with other persons in other contexts (e.g. ‘stipulative’ use in legalistic contexts).

We might imagine a similar situation affecting the Latin subjunctive, where the 1SG is unique in having a function that leads to it being interpreted as a future. One possible explanation can be found in the ‘subjunctive of determined resolution’ (cf. [Hale 1895](#), [Bennett 1898](#), [Durham 1901](#)). This type of Latin subjunctive refers to uses of the first person singular with volitive function. Thus, a phrase like *taceam nunciam* (Plautus *Bacch.*, 1058) has been interpreted to mean ‘I’m determined to keep quiet now’ ([Bennett 1898](#): 26). It has been argued that these interpretations are supported by cross-linguistic examples in Sanskrit and Ancient Greek, as proposed by [Delbrück \(1893\)](#).⁵⁰ Such examples indicate the plausibility of a reinterpretation of the subjunctive as a future, specifically in the first person singular.

In this sense, the 1SG future is not so much a form in and of itself; rather speakers simply refer to the 1SG cell of the present subjunctive when they intend to express a

⁵⁰In der ersten Person des Singular nun liegt die WILLENSERKLÄRUNG in einer Anzahl von Belegen deutlich vor. So drückt sich z. B. jemand, der sich entschlossen hat, Brahmanenschüler zu werden, und sich zu dem Zwecke der Aufnahme bei dem Lehrer meldet, so aus: *brahmacāryaśāni, ich will Brahmanenschüler werden* Çat. Br. 11, 5, 4, 1 u.ö.’ [Delbrück 1871](#): 18.

future usage. The replacement of the form in the future is thus driven by its function, which explains its limited intrusion. The other cells in the future were not replaced as their respective cells in the subjunctive did not serve similarly overlapping functions.

IMPERATIVE ALTERNATIONS (OLD CHURCH SLAVONIC): We may compare the development in Latin to the distribution of synthetic and analytic forms in imperative paradigms in Old Church Slavonic.

A clear example of a paradigm with this type of alternation can be seen in the Old Church Slavonic imperative. Table 2.24 demonstrates the forms of the imperative for the verb *ved-* ‘to lead’.

	SINGULAR	DUAL	PLURAL
1ST	—	<i>veděvě</i>	<i>veděm</i>
2ND	<i>vedi</i>	<i>veděta</i>	<i>veděte</i>
3RD	<i>vedi</i>	—	—

Table 2.24: IMP of *věděti* ‘to lead’ in Old Church Slavonic

In most person and number combinations, the imperative is formed with a synthetic form, historically from the PIE optative, formed with the ablauting suffix **-ih₁-*/**-ieh₁*. However, in the non-singular third person forms, the missing synthetic forms have been replaced with analytic constructions consisting of a combination of the particle *da* and the present indicative: *da vedq̃tъ* ‘may they lead’ (replacing **vedq̃*). The sole exception to this is the singular 3PL form *bq̃dq̃* ‘may they be’ from the verb *byti* ‘to be’.⁵¹ While the synthetic and analytic forms have different sources, the two together form the synchronic paradigm of imperatives.

The difference between synthetic and analytic forms in Old Church Slavonic is not unusual; imperative paradigms are known to alternate between synthetic and analytic forms, with the second person often presenting as synthetic, while analytic constructions

⁵¹ All other verbs have analytic forms for the 3PL.

are used to express other person and number combinations. As noted by Birjulin & Xrakovskij (2001), this lack of formal cohesion is quite common cross-linguistically. In this sense, the paradigm in Table 2.24 is just one of many.

The forms in imperative paradigms can differ widely in terms of their morphological and semantic characteristics. Imperatives are often replaced by either subjunctives (cf. Isac 2015) or with periphrastic constructions consisting of a particle plus a tensed form (cf. Slavic). With regards to their semantics, first person imperatives often do not have the same command function seen in the second person. Both French *chantons* and English ‘let’s sing!’ have more of a hortative function (Van der Auwera et al. 2003: 49-50). The third person has its own set of characteristics; consider third person constructions in English, e.g. ‘let them eat cake!’ which has a jussive connotation.⁵²

The mix of morphological and semantic characteristics puts into question whether we should consider all of these forms together as a paradigm. I argue that, at least from a diachronic perspective, it is useful to consider them as such, given the interactions between the forms. In Old Church Slavonic, the synthetic 3PL imperative *bqđq* ‘may they be’ does seem to have the same function as the 3PL of the analytic construction, e.g. *da vedqť* ‘may they lead’, despite having a distinct formal shape. Likewise, the analytic construction in all verbs besides *byti* truly seems to have replaced the corresponding synthetic forms. If true synthetic imperatives and analytic constructions did not overlap in some paradigmatic sense, there would be no motivation for the universal replacement of the synthetic form with the analytic form. Given this, I assume that we can describe these changes in terms of a paradigmatic replacement.

In summary, both the imperative in OCS and the future in Latin involve a replacement of part(s) of their paradigm by grammatical constructions that have their own

⁵²Notably, the original French version of this statement uses a subjunctive: ‘*Qu’ils mangent de la brioche*’.

syntactic structure and set of uses. This replacement is driven by overlapping function, triggering an ‘invasion’ into the local paradigm.

2.3.2.3 NON-REPLACING ALTERNATIONS: REGULAR DISTRIBUTIONS

The final two types of alternation involve a true *INVASION* by some element external to the local paradigm. This differs from the *replacement* type described in the previous section as it is both not obviously driven by usage and also does not simply copy a form from one paradigm into another. Rather, the invasive element is integrated directly into the morphology of the local paradigm.

I have distinguished examples of this type of invasion into two main categories, depending on the distribution of the invasive element. This difference is schematized in the hypothetical distributions in (3), showing a six-cell verbal paradigm:

(3) a. REGULAR DISTRIBUTION:

b. IRREGULAR DISTRIBUTION:

The first refers to cases where the invasive element has been introduced into a ‘natural’ distribution. Natural in this context refers to a grouping which can be abstracted to a coherent set of morphosyntactic properties. For example, if an invasive affix applies only to the singular, but not the plural, it obeys a natural distribution. This is schematized in (3-a).

The second type occurs when there is a disjunct distribution of forms which are difficult to motivate from a morphosyntactic perspective. For example, the distribution in (3-b) represents a distribution in which the 1SG, 2PL, and 3SG cohere in terms of

morphological expression. The environment of this distribution can only be defined in terms of a disjunct set of rules; there is simply no morphosyntactic set of features that can unify the distribution.

As we will see, the development of REGULAR DISTRIBUTIONS follow from a desire to maintain a coherent morphosyntactic distribution. That is, the distribution can be explained in terms of a natural class. As we will see, analogy (and invasion) can serve to reinforce distributions that occur along these lines. The example in the following section represents the first case in which leveling of an alternant has been motivated on the basis of a natural class. Further examples will be discussed in Chapter 4.

IRREGULAR DISTRIBUTIONS, on the other hand, do not have a straightforward explanation, and usually represent difficult patterns to account for. This type will provide the bulk of discussion in Chapter 3, while Chapter 4 discusses how analogy can interface with such a distribution.

The next few sections discuss these types in turn.

RUSSIAN AUGMENTATIVES The first type is characterized by an invasive element whose distribution can be described in terms of a natural morphosyntactic class. A clear example is the aforementioned case of the Russian diminutive in *-onok/-at-*. Here the two allomorphs are aligned with number, as shown below in Table 2.25:

	SINGULAR	PLURAL
NOMINATIVE	<i>medvež-ónok</i>	<i>medvež-át-a</i>
ACCUSATIVE	<i>medvež-ónk-a</i>	<i>medvež-át</i>
GENITIVE	<i>medvež-ónk-a</i>	<i>medvež-át</i>
DATIVE	<i>medvež-ónk-u</i>	<i>medvež-át-am</i>
INSTRUMENTAL	<i>medvež-ónk-om</i>	<i>medvež-át-ami</i>
LOCATIVE	<i>medvež-ónk-e</i>	<i>medvež-át-ax</i>

Table 2.25: *medvežónok* ‘bear cub’ in Russian

This diminutive suffix alternates between *-at-* in the plural and *-onok/-onk-* in the

singular. Despite surface differences, these forms are etymologically related (cf. Mel’čuk 1994, Gouskova & Bobaljik 2022). The picture becomes more clear if we put aside the *-(o)k* sequence from the singular.

Both *-on-* in the singular and *-at-* in the plural can trace their origins to a suffix **-ent* (cf. OCS *-et-*), which was used to make diminutives of animals (as is the case in Russian). Originally, the affix in the nominative singular *-e-* with *-et-* elsewhere.⁵³ The form in the nominative singular was extended to all forms in the singular, creating an alternation between *-e-/en-* in the singular and *-et-* in the plural.⁵⁴ Thus, analogy, specifically the extension of the nominative plural, created an alternation based around a regular distribution (singular vs plural).

At this point, the form in the singular was reinforced by the addition of the diminutive suffix *-ǫk-*. Crucially, this invasion was isolated to the singular. In other words, the coherence of a regular natural class (i.e. [+sg]) directly affected where the suffix was introduced. Its addition further solidified the distinction between the singular and plural.

A similar process occurred with other nouns in Russian. We have already seen the suppletive paradigm for ‘person’ in Russian (Table 2.6), where the singular form *čelověk* alternates with *ljud* in the plural. Another example is demonstrated by the singular *-ín*, seen below:

As with *-onok/-at*, the singular is characterized by a sequence which specifically distinguishes between the forms in the singular and those in the plural. In both cases, an intra-paradigmatic alternation falls along the lines of a natural class, producing a REGULAR DISTRIBUTION.

⁵³That is throughout all the forms in the plural as well as the singular (apart from the nominative singular).

⁵⁴The *-en-* is the result of a reanalysis of *-e-* as *-en-* underlyingly. Before vowels, the *-n-* remained.

	SINGULAR	PLURAL
NOMINATIVE	<i>sel-jan-ín</i>	<i>sel-ján-e</i>
ACCUSATIVE	<i>sel-jan-ín-a</i>	<i>sel-ján</i>
GENITIVE	<i>sel-jan-ín-a</i>	<i>sel-ján</i>
DATIVE	<i>sel-jan-ín-u</i>	<i>sel-ján-am</i>
INSTRUMENTAL	<i>sel-jan-ín-om</i>	<i>sel-ján-ami</i>
LOCATIVE	<i>sel-jan-ín-e</i>	<i>sel-ján-ax</i>

Table 2.26: *seljanín* ‘villager’ in Russian

2.3.2.4 NON-REPLACING ALTERNATIONS: IRREGULAR DISTRIBUTIONS

The second type of invasive alternation is characterized by an alternation which cannot be described in terms of a natural class. In other words, the environment has to be characterized by a set of disjunct cells. This section outlines two main examples which will subsequently form much of the discussion in the later chapters.

ROMANCE VERBS WITH -SC-: A classic example of an intra-paradigmatic alternation involving an irregular distribution is found in the present tense in several Romance languages, specifically with verbs containing reflexes of the Latin suffix -sc-.⁵⁵ Consider the following paradigms from Italian, Catalan, and Romanian in Table 2.27:

	ITALIAN ‘to finish’	CATALAN ‘to suffer’	ROMANIAN ‘to suffer’
1SG	<i>finisco</i>	<i>patéixo</i>	<i>pățesc</i>
2SG	<i>finisci</i>	<i>patéixes</i>	<i>pățesti</i>
3SG	<i>finisce</i>	<i>patéix</i>	<i>pățească</i>
1PL	<i>finiamo</i>	<i>patim</i>	<i>pățim</i>
2PL	<i>finite</i>	<i>patio</i>	<i>pățiti</i>
3PL	<i>finiscono</i>	<i>patéixen</i>	<i>pățească</i>

Table 2.27: N-PATTERN in Italian, Catalan, and Romanian (Meul 2016: 81)

The respective reflexes of the Latin inchoative suffix -sc-,⁵⁶ It. -isc-, Cat. -eix-,

⁵⁵The status of this distribution as ‘irregular’ in the sense that it cannot be explained via a morphosyntactic class is debated. Chapter 4 expands on this point.

⁵⁶The reflexes of -sc- in the modern languages are often associated with various vowels (hence Italian

and Romanian *-esc-/easc-/est-*, appear throughout the singular as well as in the 3PL, alternating with *-Ø-* in the 1PL and the 2PL.⁵⁷

The alternation between those forms with *-sc-* and those without does not have a clear synchronic motivation. Unlike the phonologically-motivated alternations discussed in Section 2.3.1.1, there is no obvious phonological reason for why *-sc-* would be deleted or otherwise left out. At the same time, there is likewise no clear syntactico-semantic reason for this distribution; the 1/2PL do not possess any special syntactic characteristics apart from the overt morphological alternation. As opposed to the Russian singulative *-onok*, the introduction of *-sc-* is not clearly based on the nature of the suffix itself. In other words, the alternation appears to be entirely morphological but at the same time cannot be explained by a natural distribution.

Chapter 4 expands on the initial introduction to the paradigm as well as the developments that affected the distribution of this suffix.

LATIN THEMATIC VOWEL IN *FERRE*: We may consider the distribution of the theme vowel in the third conjugation Latin verb *ferō* as another potential example of an invasive alternation. The present indicative of *ferō* is characterized by a mix of forms with a theme vowel and those without:

	SINGULAR	PLURAL
1ST	<i>fer-ō</i>	<i>fer-i-mus</i>
2ND	<i>fer-s</i>	<i>fer-tis</i>
3RD	<i>fer-t</i>	<i>fer-u-nt</i>

Table 2.28: PRES IND of *fer-* ‘to bear’ in Latin

In the 2SG, 2PL, and 3SG, the ending is added directly to the root. However, the

-isc-).

⁵⁷An alternative characterization of the alternation is that the suffix alternates with a theme vowel *-i-* in the first and second plural. Italian complicates this picture, as the ending *-iamo* has been generalized to all verb classes from *stiamo* ‘we stand’ (Meyer-Lübke 1894: 167). However, the general alternation is clear; the reflexes of *-sc-* have been appended to a theme vowel that varies depending on person and number.

1PL and 3PL are characterized by a vowel in between the root and the ending. The 1SG on the surface appears to append the ending directly to the root; however, from a historical perspective, there are reasons for associating this form with those that have an extra vowel, as will be expanded on in Chapter 3.

Thus we are left with an alternation in which a *-i-* and *-u-* appear in the 1PL and 3PL, while the other forms have no vowel at all. This alternation contrasts with other third conjugation verbs, as we can see below in Table 2.29:

	FERŌ ‘I bear’	GERŌ ‘I bear’
1SG	<i>fer-ō</i>	<i>ger-ō</i>
2SG	<i>fer-s</i>	<i>ger-i-s</i>
3SG	<i>fer-t</i>	<i>ger-i-t</i>
1PL	<i>fer-i-mus</i>	<i>ger-i-mus</i>
2PL	<i>fer-tis</i>	<i>ger-i-tis</i>
3PL	<i>fer-u-nt</i>	<i>ger-u-nt</i>

Table 2.29: PRES IND of *ferō* ‘I bear’ and *gerō* ‘I bear’ in Latin

The comparison between *ferō* and *gerō* demonstrates the lack of thematic vowel in the 2SG, 2PL, and 3SG clearly. The alternation can thus be described as one between thematic and athematic forms.

Chapter 3 will expand on why we find both athematic and thematic forms in the paradigm of not only this verb in particular, but a disparate set of irregular verbs. Ultimately, we will see how a disparate set of irregular verbs has conditioned the appearance of athematic and thematic forms in a disjunct set of environments.

What unites Latin and Romance examples is that the distributions of alternants cannot be described with a natural class of morphosyntactic features. It is clear that we cannot appeal to a paradigmatic replacement or an invasion on the grounds of the nature of the suffixes alone. I have thus not given an explanation for how these affixes entered their respective paradigms and why they have such peculiar distributions. Section 2.4 expands on a previous description of a phenomenon that makes reference to this

type of distribution, particularly in the works of Maiden, who calls these distributions *morphomes* (Maiden, 2003, 2005, 2018, *etc.*). Chapter 3 and Chapter 4 will further examine the Romance and Latin examples respectively.

2.3.3 COMPLICATIONS: CONFLATION AND LEVELING

Before summarizing this section, it is worth briefly noting that membership in one of these classifications is not set in stone. As we have already seen, specific paradigms can exhibit several different types of alternation. In particular, alternations which were originally created through analogical change can intersect with phonological changes that create new competing forms.⁵⁸ All these changes can obscure relatively straightforward diachronic developments to create a set of fairly disparate alternants. In this sense, it is possible for different types of alternations to apply to the same paradigm.

Another factor that can complicate the diachronic origin of alternations is analogical leveling or extension. That is, the original distributions of alternants are themselves susceptible to change. Take, for example, an irregular distribution like that in (3-b). We might imagine various possible developments where the distribution changes based on some process of analogical leveling or extension. Two possible pathways for analogical change to illustrate this can be schematized as below:

(4)

a.

INITIAL DIST:

b.

OPTION A:

c.

OPTION B:

We may observe either OPTION A or OPTION B without realizing that the original

⁵⁸See the Romanian example in Section 2.3.1.2 for an illustration.

distribution was different. We must therefore recognize that it is possible that leveling or extension have affected originally irregular patterns. Chapter 4 looks into this type of situation in more depth, examining what factors affect the directionality of change.

2.3.4 INTERIM SUMMARY

The last few sections have expanded on the different types of intra-paradigmatic alternations expressing non-root morphology, repeated below:

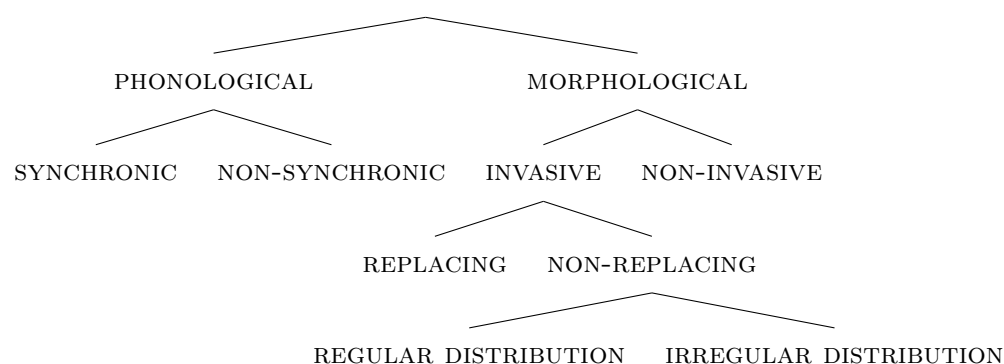


Figure 2.2: Types of non-root alternations

To summarize, the main distinction is based on whether the origin of the alternants is phonological or morphological in nature. These types can be further divided, depending on various criteria like their synchronic accessibility and whether the alternants are internal or external to the paradigm in origin.

2.4 MORPHOMIC DISTRIBUTIONS

The distributions described in Section 2.3.2.4 represent highly idiosyncratic morphological patterns. However, such distributions are not limited to non-root intra-paradigmatic alternations. In a long-running research program, Maiden (2003, 2004, 2005, 2018) examines irregular distributions in Romance which encompass not only non-root phe-

nomena but also patterns of suppletion and stem alternations involving phonological changes like diphthongization and apophony. These phenomena seem to cohere together in terms of their morphological exponence, a surprising fact given that we might not expect them being distinct types of morphological exponence. Maiden calls these patterns *morphomes*, a term which is crucially not to be confused with *morphemes*. This section describes morphomes in greater detail and discusses the nature of these patterns as they pertain to diachronic and synchronic perspectives.

2.4.1 THE TERM ‘MORPHOME’

The term ‘*morphome*’ was first used by Aronoff (1994) to describe phenomena that are purely morphological in nature and not irreducible to either syntax or phonology. Aronoff illustrates this concept through the English ‘perfect participle’ (cf. Aronoff 1994: 23-25). This participle has two separate syntactic functions but is identical in form in each (examples cited from Maiden 2018: 10):

- (5) Passive: *It is seen/written/sought/read/torn/brought/taken*
 Perfect: *I have seen/written/sought/read/torn/brought/taken*

The purpose of Aronoff’s illustration of the English perfect participle is to describe an aspect of grammar that cannot be characterized in either morphosyntactic terms, or in what he calls ‘morphosyntactic feature arrays’ (Aronoff 2016: 24), or in phonological terms. From a morphosyntactic perspective, the uses of the participle in (5) cannot be described with one feature matrix, given that the passive use implies a [+passive] feature, while the perfect use requires some sort of perfective feature, e.g. [+perfect]. Crucially, the passive use in (5) is not perfect and vice versa. The morphosyntax therefore requires two morphosyntactic features at two different instances. Aronoff therefore supposes that

the exponence in these two contexts cannot be explained by a unified morphosyntactic characterization.

The syncretism of the two constructions likewise cannot be explained in phonological terms. The participial ending has several allomorphs whose distribution is not solely determinable by phonology. The most productive allomorph is *-D* in Aronoff’s terms (Aronoff 1994: 23), which comes in three shapes, phonologically conditioned by the final segment of the verb stem: /-d/ in *loved* /lʌv-d/, /-t/ in *baked* /beɪk-t/, and /-əd/ in *lifted* /lɪft-əd/. Another allomorph is the suffix *-en*, as in *written*. These suffixes may also be accompanied by stem changes; for example, the vowel is changed as in the pair *keep-kept*, while the entire stem is changed in the pair *catch-caught*. While the distribution of the allomorphs of the first suffix *-D* can be captured in phonological terms, the distribution of the other allomorphs is synchronically arbitrary, suggesting that it is not determined by the phonological component of the grammar.

As the notion of morphemes as anything which cannot be described via phonological or morphosyntactic means is quite broad. Aronoff (2016) distinguishes between several types. However, the use of the term morpheme has been further used widely in the literature to refer specifically to what Aronoff (2016) calls a ‘polyvalent polymorphous mapping’.⁵⁹ This term refers to cases where a paradigmatic set of disjunct cells have an identical form, despite there not being a syntactic basis for the similarity in exponence.⁶⁰ The use of morpheme to describe this specific phenomena has been particularly widespread in the literature on Romance morphology, following work by Maiden (e.g. Maiden 2018). It is this notion of morpheme which is relevant for our purposes, as it applies to cases of non-root alternations with irregular distributions.

⁵⁹For more information see Aronoff (2016: 24-25).

⁶⁰This means that accounts that try to refute morphemicity often propose underlying structural similarities between the supposed disjunct cells.

2.4.2 MORPHOMES AS TEMPLATES

The definition of morphome that I adopt is based on Maiden’s work on Romance. For Maiden, a morphome is an abstract distribution that can be placed upon various morphological phenomena to produce a pattern which is reflected on the surface via coherence in form.

To illustrate this, consider the distribution of the Italian verbs *andàre* ‘to go’ and *finìre* ‘to finish’ from Table 1.1, reproduced below:

	ANDÀRE ‘to go’	FINÌRE ‘to finish’
1SG	<i>vàd-o</i>	<i>finisc-o</i>
2SG	<i>và-i</i>	<i>finisc-i</i>
3SG	<i>và</i>	<i>finisc-e</i>
1PL	<i>and-iàmo</i>	<i>fin-iàmo</i>
2PL	<i>and-àte</i>	<i>fin-ite</i>
3PL	<i>vàn-no</i>	<i>finisc-ono</i>

Table 2.30: PRES IND of *andàre* ‘to go’ and *finìre* ‘to finish’ in Italian

Both paradigms represent distinct types of morphological phenomena, i.e. i) root suppletion, and ii) suffixation of *-isc-*. However, the intra-paradigmatic distributions are generalizable across both paradigms respectively. The cells where there is a *va(d)-* allomorph correspond exactly to the cells where *-isc-* is present. This pattern not only encompasses these two examples, but also governs several classes of verbs through patterns of suppletion, suffixation, and apophony. I expand on these examples in Chapter 4. Simply put, there exists an abstract distribution contrasting two sets of cells, i) the singular and third plural, and ii) the first and second plural, schematized below:

Table 2.31: Morphomic template (Italian)

For our purposes, a morphome is therefore a morphological abstraction which con-

ditions the exponence of the cells it targets. Its abstract nature is evidenced by the fact that it targets different types of morphological expression (e.g. suppletion, suffixation, etc.), with the patterns maintaining a morphological coherence within the specified templatic domain.

2.4.2.1 CRITERIA FOR MORPHOMEHOOD

Having established what a morpheme is, it is necessary to define what criteria are required to establish a morphomic template, as this section will now aim to do. This section describes a set of possible criteria.

Maiden (2018: 19-22) lays out two criteria that are necessary for morphomicity. The first is that the distribution of the cells affected by the morphomic template is irreducible to either phonological or syntactic factors (or *grammatical features* in the words of **Maiden 2018: 20**).⁶¹ This criterion establishes that the distribution must be explained via a morphological component, thus satisfying the original notion of a morpheme as proposed by **Aronoff (1994)**. In practice, irregular distributions such as those described in Section 2.3.2.4 satisfy this criterion by definition, as they cannot be referenced by a single morphosyntactic set of features.

The second criterion requires there to be a ‘generalization’ across a set of unrelated phenomena to be present. A generalization in this context refers to patterns in which multiple types of morphological phenomena cohere in their distribution. For the suppletive alternation in *andare* ‘to go’, the alternants are set up in a generalizable distribution; *va(d)* is found in the singular and third plural, while *and* is found in the first and second plural. What is crucial for morphomicity is that this distribution must have morphological coherence across several types of morphophonological phenomena.

⁶¹Of course, it is possible that some (or all) of the morphomic distributions can actually be described in terms of morphosyntactic feature sets, depending on how abstract the framework is (cf. **Zompi 2023**).

As Maiden (2018: 20) states, ‘this latter consideration (“the same ... but with different realization”) is precisely what guarantees that such data are morphomic: it shows that the pattern is general for more than one type of alternant’. What Maiden is referring to here is the fact that morphomic patterns are not only found with one type of morphological phenomenon, but rather with several different types of alternants. For example, the Romance case demonstrates suppletive, phonological, and suffixal alternations. It is precisely the fact that all of these phenomena show morphological coherence across the same set of cells which supports Maiden’s conclusion that this template constitutes a psychologically real object in the grammar.

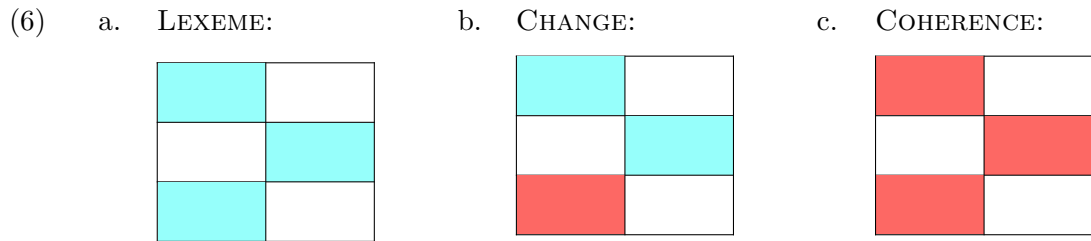
In sum, we have two criteria which we can use to determine morphomicity, i) a distribution cannot be explained either phonologically or (morpho)-syntactically, and ii) a distribution must show coherence across several types of morphological expression (e.g. suppletion, affixation, periphrasis, apophony, etc.).

2.4.2.2 MORPHOMES IN DIACHRONY: DIAGNOSTICS AND ANALOGY

Morphomes in Maiden’s view are inextricably linked to diachrony. The key evidence for the existence of morphomes as psychological real objects which are accessible to speakers is based on diachrony (Maiden, 2018). Thus, it is important to briefly discuss the interaction between morphomic patterns and diachrony. There are three types of diachronic changes which Maiden uses to evidence the psychological reality of morphomes, i) coherence, ii) convergence, and iii) attraction.⁶²

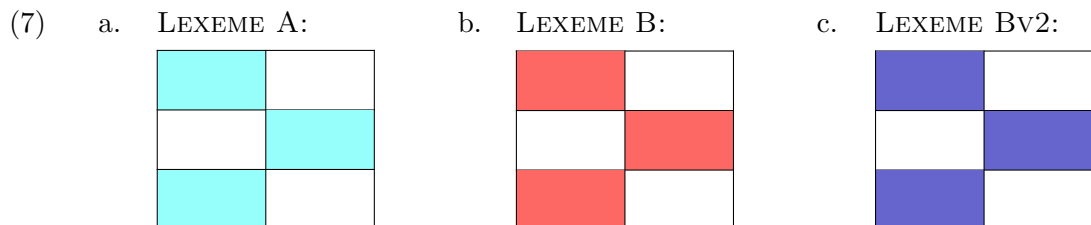
COHERENCE: Coherence refers to the notion that all the cells that compose the morphome are coherent in form across the board. If any change affects one of the cells, it will also spread to the other cells. Consider the schema in (6).

⁶²See the discussion of these diagnostics in Maiden (2018: 10-17).



Let us assume that a lexeme has some stem alternation which follows the morphomic distribution of the cells shaded in blue in (6-a). Suppose some innovation introduced a change to a paradigm, represented with the shading in red as in (6-b). This distribution would not be coherent, as the cells are not uniform in their form. If, however, all the cells change, as seen in (6-c), it is evidence that the distribution represents a psychologically real object.

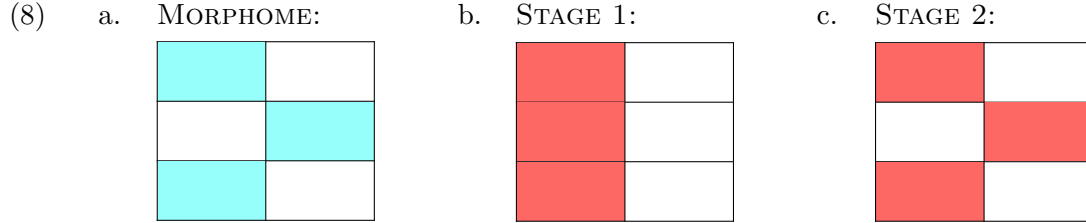
CONVERGENCE: Convergence occurs when a phonological property of a morphomic lexeme is generalized to other lexemes which participate in the morphomic distribution. Consider a new schema as in (7):



Let us assume we have two lexemes, shown in (7-a) and (7-b), which both engage in the morphomic distribution. Lexeme A has some phonological characteristic represented with the blue shading. Lexeme B lacks this characteristic, and is thus shaded with red. If Lexeme B takes on the phonological property associated with Lexeme A, as shown in (7-c), the two lexemes have *converged* with respect to their form. Crucially, all the cells

of the morphome engage in the convergence.⁶³

ATTRACTION: The final diachronic change, attraction, occurs when a morphomic pattern is applied to new lexemes, forcing allomorphs to be organized along the templatic lines. Such an attraction can be schematized as below:



Let us assume that there exists a morphomic distribution as represented by the cells shaded in blue in (8-a). This distribution applies to several lexemes involving different types of phenomena. Consider then a new lexeme, with a morphological distribution seen in (8-b). This distribution is distinct from the morphomic distribution. Now consider the next diachronic step, in which the distribution of the new lexeme has changed to cohere with other words in the language. This new distribution is represented in (8-c). What has happened here is that the morphomic distribution has *attracted* the new lexeme to fit into the distribution in (8-a).

MORPHOMES AND PROPORTIONAL CHANGE How do these types of changes fit with other types of morphological change? In many ways, each of these changes seem to be describable in terms of an analogical proportion, albeit one that references an abstract proportion encompassing a set of forms rather than one specific form. It is clear that the changes discussed above are in principle due to analogy to other lexical items.

⁶³This type does not play a role in further discussion in this dissertation. An example of this change in Romance occurs in the PYTA morphome (forms which are descended from the Latin *perfectum*, cf. Maiden 2018) in which high vowels spread from the 1SG preterite to other cells in the morphome across the board.

However, it is difficult to describe this type of change via a specific proportion, precisely because of the abstract nature of the comparison. Let us consider morphomic attraction as an example. Firstly, when morphomic attraction occurs, it does not involve the comparison between two cells; rather it can involve multiple cells which cannot even be described coherently without reference to an arbitrary template. Secondly, the actual morphological exponence of the attraction depends on the items themselves. In the proportion *stone* : *stone-s* :: *cow* : *X*, *X* = *cow-s*, the comparison is direct as the suffix *-s* is being added to both forms. For morphomic attraction, the allomorphs involved between lexeme *A* and lexeme *B* might be describable in terms of completely distinct morphological phenomena, e.g. suppletion and periphrasis. In this sense, this morphological change, while fairly intuitive, is difficult to describe in traditional terms.⁶⁴

MORPHOMES: PRODUCTIVITY AND LEARNING MODELS If we consider the extension of morphomic patterns in terms of productivity, we find a similar descriptive challenge. Ultimately, the reasons for the difficulty in describing these changes using traditional proportional analogy also apply to analyses involving productivity.

Let us consider once again the pattern seen in Italian, where *andàre* and *finìre* both alternate in form (see Table 2.30). The distributions are identical; however, the actual surface alternants could not be more distinct: *and~va(d)* and *fin~finisc*. If we had to account for a single surface rule to unify these verbs, it would be impossible. What unites the two verbs is an abstract distribution, not their surface form. Both the Tolerance Principle and the MINIMAL GENERALIZATION LEARNER (cf. Section 1.4) are dependent on a direct comparison between surface forms in order to make and learn rules. Given that the comparison between the verbs that engage in the morphome are not based on what is seen on the surface, morphomes prove challenging for these learning algorithms

⁶⁴Similar conclusions can be drawn about the other types of change discussed above.

to model. For the MGL, there exists a base that can derive each cell in the present tense of *finire*. However, the surface differences between *andare* and *finire* makes it difficult for that base to also derive all forms in *andare*. The forms in the 1/2PL in *andare* are not derivable by the same means as *finisc* and *fin*.⁶⁵

However, this is not to say that it is impossible for these models to account for these distributions. For the Tolerance Principle, we need only assume that the rule that is being learned is abstract and is not based on the exact surface strings. To prefigure the discussion in Chapter 3, Embick (2016) accounts for the Italian data with an abstract impoverishment rule. If it is this impoverishment rule that is deemed productive by the Tolerance Principle, it may be extended to novel verbs. In this sense, the intuition behind the Tolerance Principle can still explain morphomic distributions as well as their extension to novel forms.

2.4.2.3 SUMMARY: MORPHOMES

This section has discussed the nature of morphomes. On the basis of Maiden (2018), I have set up two criteria for diagnosing morphomicity, including whether there is an irregular distribution and whether different morphological phenomena engage in the distribution. I have also discussed morphomes from a diachronic perspective. What is clear is that morphomes have a descriptive reality in terms of diachronic change. The fact that *andare* and *finire* have the same distribution, for example, simply cannot be a coincidence.⁶⁶ However, as discussed in this section, morphomes pose a challenge for both traditional notions of analogical change as well as for recent models based on productivity and learning models, specifically because of the abstract nature of the

⁶⁵It might be possible to make reference to *principle parts* in order to provide multiple bases which can be used to describe the present tense using the MGL.

⁶⁶We will also see other examples in the subsequent chapters.

comparisons involved.

2.5 ALTERNATIONS AND SYNCHRONY

The general notion of non-root alternations is not particularly difficult to describe from a synchronic perspective. Phenomena like affixation and periphrasis are not rare and are thus describable in most if not all morphological frameworks. However, this is not to say that studying their distributions is uninformative for synchronic theory. This section discusses some of the problems that closer examination of these alternations can give insight into. I focus on two specific issues, i) the status of stems, and ii) the status of morphemes, both from a synchronic distribution.

2.5.1 STATUS OF STEMS

One big divide between different morphological frameworks pertains to the status of stems in the theory. On the one side, some theories (cf. DISTRIBUTED MORPHOLOGY Halle & Marantz 1993) deny the existence of stems as having a special status (Embick & Halle, 2005). Rather, stems can be decomposed into smaller pieces involving the root combining with individual affixes. On the other side is the position that stems do hold a special status as morphological objects, as argued by Aronoff (1994), *a.o.* This has led some morphological frameworks, such as PARADIGM FUNCTION MORPHOLOGY (Stump, 2001, 2015) to have special morphological functions that make reference to stems. The debate is still ongoing, with recent work by Kodner (2022) arguing in favor of a structural interpretation of stems via syntactic spans.

Affixal alternations naturally fit in this discussion, as they often make up part of what could be considered a stem. For example, we might describe the alternation in Italian *finire* to be between *finisc-* and *fin-*; the alternation would then be between two

stems instead of two affixes, making a direct comparison with *andare* more reasonable as we can also describe the difference between *va(d)-* and *and-* as stem-based. On the other hand, describing the alternation this way seems to obfuscate the fact that *-isc* independently attaches to other verbs which also engage in the same alternation (e.g. *capire* ‘to understand’ alternates between *capisc-* and *cap-*). Describing the suffix *-isc* as an independent morpheme is the most straightforward way to generalize over the class of verbs which show this alternation. These two perspectives demonstrate the connection between affixal alternations and the discussion of the status of stems as indivisible units that are able to be referenced by the grammar.⁶⁷

In Chapter 3, I discuss this issue in more depth, examining the alternation affecting the Latin verb *ferre* (see Table 2.28). By comparing two potential analyses, one in DISTRIBUTED MORPHOLOGY (DM) and one in PARADIGM FUNCTION MORPHOLOGY, I explicitly contrast the two sides of the issue, as one of the differences of the two frameworks is their treatment of stems. I ultimately favor a piece-based analysis in DM, showing that the abstract morphomic patterns in the Latin verb *ferō* as well as a few other irregular verbs are best explained in terms of a stem-less approach.

2.5.2 STATUS OF MORPHOMES

While the evidence for morphomes centers around their diachronic patterns of attraction, their status has been a topic of considerable controversy in the synchronic literature. On the one hand, some morphological theories readily accept their existence, implementing specific mechanics that make reference to morphomes (e.g. *morphomic properties* in PFM). Other models adapt morphomes into already established principles (cf. Blevins 2006, 2016, Blevins et al. 2019). On the other hand, theories like DM take

⁶⁷I subsume cases where stems exist but are divisible under non-stem approaches for this purpose.

morphomic patterns to be epiphenomenal (e.g. Embick 2016, Trommer 2016). Rather than specifying that morphemes reflect distinct entities, rules underlie the irregular distributions, causing changes in exponence that are independent from the distributions themselves. Given that affixal and periphrastic alternations are sometimes attracted to morphomic distributions, these theories must also be able to account for their behavior.

In Chapter 3 and Chapter 4, I adopt the latter position, namely that morphomic distributions are epiphenomenal. Building on the work of Embick (2016), in which the Italian distribution is explained via an impoverishment rule which forces the first and second plural to express default morphology, I examine both the Romance and Latin distributions discussed in Section 2.3.2.4. In Latin, I explicitly compare DM and PFM, demonstrating that a *morpheme*-based approach is superior in terms of accounting for the meaningful abstract generalizations underlying the putative morpheme in comparison to an approach which makes specific reference to morphemes as entities. As we will see, this argumentation challenges Maiden’s notion of morphomic coherence, as the verbs which partake in the distribution do not show the same morphomic cohesion, despite being underpinned by a meaningful morphological generalization in terms of athematicity.

For the Romance case, I show that changes in the distribution, via the leveling of some of the alternants, do not support a templatic approach. Overall, I argue that morphomic (and thus irregular) distributions can be decomposed into disjunct environments, which are better accounted for with underlying impoverishment rules which subsequently affect surface forms, albeit indirectly.

2.6 CONCLUSION

This chapter has aimed to provide a foundation for subsequent discussion in Chapter 3 and Chapter 4. To begin with, I have provided a system by which non-root intra-paradigmatic alternations can be classified, in accordance with their diachronic origin. As the examples show, these alternations consist of a wide range of phenomena. Yet, generalizations about their diachronic and synchronic behavior can be made.

The next chapters discuss two case studies of irregular, or *morphomic*, alternations. As we have seen, the distributions of morphomic patterns are highly idiosyncratic and thus demand sophisticated diachronic and synchronic explanations. In this sense, irregular distributions may provide insight into the study of both diachrony and synchrony. With respect to the former, Chapter 3 discusses the origin of a distinctive pattern of athematicity in Latin, which coincides with the development of a periphrastic alternation. Chapter 4 moves on to examine how leveling can disrupt morphomic distributions, an aspect which has been neglected.

From a synchronic perspective, the subsequent chapters examine morphomic patterns insofar as they relate to the following themes: i) the status of stems, and ii) the status of morphemes in synchronic theory.

perīre mē mālīs mālīm modīs.

Plautus, *Bacchides* 490

3

Synthetic-Analytic Alternations

3.1 INTRODUCTION

As we saw in the last chapter, non-root alternations come in many forms, with several diachronic sources. However, there remain examples whose distribution is still unclear from a diachronic perspective. In many such cases, one can appeal to an analogical relationship with other verbs that demonstrate morphological alternations in the same distribution (cf. *morphomic convergence*). In what follows, we will examine a potential example of such a development.

This chapter discusses the origin of a periphrastic alternation with an irregular distribution. As we will see shortly, the Latin verbs *nōlō* ‘I do not want’ and *mālō* ‘I prefer’ demonstrate highly idiosyncratic patterns which have hitherto been overlooked from both diachronic and synchronic perspectives. Despite their irregularity, these verbs engage in a regular alternation of athematicity, exhibited by several other verbs. Together, these distributions inform our understanding of i) how analogy with abstract morphomic patterns can contribute to the formation of periphrasis (a type of non-root alternation), and ii) how these distributions are encoded from a synchronic perspective.

The structure of this chapter is as follows. Section 3.2 outlines the facts of the distribution of the various relevant verbs in two distinct stages of the attested Latin period, the Classical period and the Old Latin period. This includes the observation that many irregular verbs in Latin engage in an irregular athematic pattern. Secondly, I give a novel explanation for the distribution of the uncontracted variants of *nōlō* ‘I do not want’ and *mālō* in Plautus, based on the position in the verse. This section concludes by providing the relevant diachronic background of these verbs, giving context for the discussion in Section 3.3.

Section 3.3 summarizes the issue of the diachronic origin of the periphrasis and discusses the benefits and drawbacks of different approaches. Later in the section, I provide two potential solutions which may have influenced the creation of the periphrastic alternation. This diachronic discussion segues into Section 3.4, which examines the morphomic distribution in Classical Latin from a synchronic perspective. Here, I compare two contemporary frameworks, DISTRIBUTED MORPHOLOGY (Halle & Marantz, 1993) and PARADIGM FUNCTION MORPHOLOGY 2 (Stump, 2015), with respect to the phenomenon of athematicity. In doing so, I demonstrate that PFM2 fails to provide a satisfactory account for the distribution. Section 3.5 concludes.

3.2 ESTABLISHING THE FACTS

The present indicative paradigms of the verb *volō* ‘I want’ and its derivatives *nōlō* ‘I do not want’ and *mālō* ‘I prefer’ are characterized by their idiosyncratic alternations involving periphrasis and contraction, shown below in Table 3.1.

	VOLŌ ‘I want’	NŌLŌ ‘I do not want’	MĀLŌ ‘I prefer’
1SG	<i>volō</i>	<i>nōlō</i>	<i>mālō</i>
2SG	<i>vīs</i>	<i>nōn vīs</i>	<i>māvīs</i>
3SG	<i>vult</i>	<i>nōn vult</i>	<i>māvult</i>
1PL	<i>volumus</i>	<i>nolumus</i>	<i>mālumus</i>
2PL	<i>vultis</i>	<i>nōn vultis</i>	<i>māvultis</i>
3PL	<i>volunt</i>	<i>nōlunt</i>	<i>mālunt</i>

Table 3.1: PRES IND of VOLŌ, NŌLŌ, and MĀLŌ in Latin

The verb *volō*, like most verbs in Latin, has a ‘synthetic’ present indicative, in which the verbal suffixes expressing tense/mood, and person marking are added directly to the root as one word.¹ However, its derivatives diverge from other verbs by exhibiting ‘analytic’ behavior in the present indicative where other verbs have synthetic behavior.

To exemplify the unique behavior, consider the first and third singulars of *volō* and *nōlō* as seen in Table 3.1. In the first singular, both verbs have a ‘synthetic’ form, where the first person singular suffix *-ō* is added to the stems *vol-* and *nōl-*. In the third singular of *volō*, the suffix *-t* combines with the root *vol-* to produce *vult*. In terms of synthesis, both the first singular and third singular of this verb pattern with one another.

The third singular of *nōlō*, however, differs from the first singular; instead of a synthetic form, one finds an ‘analytic’ combination of the third singular form of *volō* (i.e. *vult*) the negative particle *nōn* ‘not’, producing an alternation between the first and

¹This is not to say that the verb does not exhibit irregularity in this paradigm, as we will see shortly.

third singular in terms of periphrasis. As we can see in Table 3.1, analytic forms are also present in the second singular (*nōn vīs*) and the second plural (*nōn vultis*). The present indicative of *nōlō* can thus be said to exhibit a ‘synthetic-analytic’ alternation.

The second derivative of *volō*, namely *mālō*, shows a slightly different but nonetheless similar alternation in the present indicative. The base form of the verb is *māl-* which itself is a contraction of *mā-* and *vol-*. However, this root has an alternate *māvul-*, where the contraction between *mā-* and *vol-* has seemingly not occurred.

If we once again focus on the first singular and the third singular, we can illustrate this difference between the forms *mālō* and *māvult*. The 1SG simply adds the person ending *-ō*, exactly paralleling *volō* and *nōlō*. The third singular, however, adds *mā-* to the third singular of *volō*, producing *māvult* from *mā+vult*. There is thus an alternation between a ‘contracted’ form of the root involving the sequence *māl-* and an ‘uncontracted’ form of the root involving the prefixation of a sequence *mā-*. Unlike *nōlō* all forms in the present indicative are ‘synthetic’; there is no periphrasis in any of the forms; the alternation is expressed in terms of contraction and non-contraction. As we can observe by comparing the forms of *nōlō* and *mālō* in Table 3.1, the distribution is notably identical between the two verbs. There is periphrasis and non-contraction in the second person singular and plural as well as the third person singular.

The following section outlines the synchronic distribution of these forms and discusses various idiosyncratic properties of each verb as found in two distinct stages of Latin, that of the early period and that of the Classical period. Section 3.2.1 begins by discussing the general properties of the verbs in question during the Classical period, as the distribution of forms at this stage is well-established. This segues into a comprehensive investigation of the earlier state of *volō* and its derivatives as found in the plays of

Plautus, a playwright in the early period of Latin.² As we will see, the patterns found in these verbs during both of these stages is related to a broader pattern of athematicity and suppletion seen within a small class of Latin irregular verbs. Section 3.2.4 will end the discussion of the synchronic properties by discussing these irregular verbs and how their properties relate to the discussion of the alternations in the present indicative of *nōlō* ‘I do not want’ and *mālō* ‘I prefer’.

3.2.0.1 SYNTHETIC-ANALYTIC VERSUS PERIPHRAISIS

Before continuing to lay out the synchronic properties of each verb, it is necessary to define some of the terminology used in this chapter. Firstly, I broadly refer to alternations as seen in Table 3.1 as synthetic-analytic alternations. We can describe the 1SG *nōlō* as synthetic and the 3SG *nōn vult* as analytic. In the literature, the analytic forms in the paradigm are often referred to as periphrastic. For the purposes of cohesion, I use the terms SYNTHETIC and ANALYTIC in this chapter as well as elsewhere.

A second difficulty arises when discussing the forms of *mālō*. Even though the idiosyncratic alternation between forms with a contracted *māl-* and an uncontracted *mā-* prefix are found with the same distribution as the ANALYTIC forms of *nōlō*, the alternation is strictly not periphrastic. Rather, the alternation is expressed in terms of (non-)contraction; every form in the paradigm is actually synthetic. Thus, the parallel between *nōlō* and *mālō* cannot be described simply in terms of a SYNTHETIC-ANALYTIC alternation. For this reason, I introduce the term SEMI-ANALYTIC to describe the forms of *mālō* which do not fully contract. Thus *māvult* is SEMI-ANALYTIC while *mālō* is SYNTHETIC.

To summarize, there are three different types of forms at play. The term SYNTHETIC

²See Section 3.2.2.

refers to forms which are one word and, if they involve a combination of several elements, are fully contracted, e.g. *nōlō* and *mālō*. The term ANALYTIC refers to forms which are made up of more than one word, e.g. *nōn vult*. Finally, the term SEMI-ANALYTIC refers to forms that are partly contracted and thus are one word, but are still segmentable into two separate parts, e.g. *māvult*.

3.2.1 SYNCHRONIC DISTRIBUTION IN CLASSICAL LATIN

The term Classical Latin refers to the period of Latin spoken and written from the first century BCE up to the third and fourth centuries CE. The texts and grammar of this period make up the usual content of the Latin that most people are familiar with and represent the basis for most of the synchronic discussion of the Latin language. This section focuses on the distributions of *volō* and its two derivatives *nōlō* and *mālō* during this period, during which the forms of these verbs show uniformity.³ The next section compares the Classical distribution with that found in Old Latin, focusing specifically on the texts of Plautus.⁴

3.2.1.1 VOLŌ IN CLASSICAL LATIN

The present indicative of the irregular verb *volō* ‘I want’ is peculiar with respect to three characteristics: i) root allomorphy in the second singular, ii) athematicity in the second person and third singular forms, and iii) a thematic vowel with irregular vowel quality in the first plural. To illustrate these characteristics, consider the differences between the third conjugation verb *colō* ‘I cultivate’ and *volō* ‘I want’ in Table 3.2.

³See Appendix A.1 for the full paradigms of *volō*, *nōlō*, and *mālō*.

⁴For more information on the different periods of Latin, see Weiss (2020: 24-25). For our purposes we are concerned with Old Latin (3rd-2nd century BCE) and Classical Latin as outlined above. The texts in the period before Old Latin during the Very Old Latin period (6th-3rd century BCE) do not offer much information that sheds light on the distribution of these forms. For this reason, I focus the discussion by comparing the distribution of the Classical and Old Latin periods.

	COLŌ ‘I cultivate’	VOLŌ ‘I want’
1SG	<i>col-ō</i>	<i>vol-ō</i>
2SG	<i>col-i-s</i>	<i>vī-s</i>
3SG	<i>col-i-t</i>	<i>vul-t</i>
1PL	<i>col-i-mus</i>	<i>vol-u-mus</i>
2PL	<i>col-i-tis</i>	<i>vul-tis</i>
3PL	<i>col-u-nt</i>	<i>vol-u-nt</i>

Table 3.2: PRES IND of *colō* ‘I cultivate’ and *volō* ‘I want’ in Latin⁵

Despite having nearly homophonous roots, *volō* is distinguished from *colō* in terms of the three aforementioned characteristics. The second singular *vīs* exhibits irregular root allomorphy, in contrast with the regular *colis*. The second person (*vīs* and *vultis*) and third person singular (*vult*) forms are *athematic*; that is, they lack a theme vowel, found overtly in the paradigm for *colō* as *-i-* (*colis*, *colit*, *colimus*, *colitis*) or *-u-* (*colunt*).⁶ Even where a theme vowel is present, it does not necessarily align with the regular outcome; in the first plural the theme vowel is *-u-* (*volumus*) instead of *-i-* (*colimus*). Besides these irregularities, the present indicative of *volō* behaves like a third conjugation verb like *colō*.⁷

⁵The segmentation in these forms is my own. I will justify this segmentation in comparison with alternative approaches in Section 3.4.

⁶While the first singular does not have an overt theme vowel, I assume that a theme vowel is present underlyingly. There are a few facts that support this assumption. In the second and fourth conjugations, the thematic vowel specific to these conjugations is present in the first singular: *-eō* from *-ēō* and *-iō* from *-īō*. An underlying short vowel would be deleted before the long *-ō*, meaning that for both *colō* and *volō* the first singular underlyingly could very well have been *-eō* → *-ō*. Furthermore, the ending *-ō* was diachronically the thematic ending for the first person singular present indicative. A reflex of the athematic ending **-mi* is not present in these forms, suggesting that over the course of the development of the different conjugations of Latin, the first singular was analyzed as thematic.

⁷The verb *volō* does display irregularity in its other tenses and moods. It has an irregular present subjunctive, mirroring the irregular present subjunctive of *esse* ‘to be’. As we will see shortly, its present infinitive and imperfect subjunctive also show a lack of a theme vowel, mirroring some of the forms of the present indicative. Finally, the verb notably lacks a passive and imperative. These idiosyncrasies are inherited by its Latin derivatives *nōlō* and *mālō*. A peculiarity is the unique negative imperative construction formed with the subjunctive stem *nōlī*: e.g. 2PL *nōlīte* ‘don’t want’; the verb *volō* has no imperative.

The verb *volō* has largely been regularized in the modern Romance languages. Few traces remain of *nōlō* and *mālō*.

Interestingly, the scope of athematicity in the paradigm of *volō* extends beyond the present indicative. Both the imperfect subjunctive as well as the present infinitive show forms lacking a theme vowel. For comparison, consider the present infinitive (Table 3.3) and the imperfect subjunctive (Table 3.4) for *volō* ‘I want’ and *colō* ‘I cultivate’.

VOLO	<i>vel-le</i>	‘to want’
COLO	<i>col-e-re</i>	‘to cultivate’

Table 3.3: PRES INF of *volō* and *colō* in Latin

	1SG	2SG	3SG	1PL	2PL	3PL
VOLO	<i>vellem</i>	<i>vellēs</i>	<i>vellet</i>	<i>vellēmus</i>	<i>vellētis</i>	<i>vellent</i>
COLO	<i>colerem</i>	<i>colerēs</i>	<i>coleret</i>	<i>colerēmus</i>	<i>colerētis</i>	<i>colerent</i>

Table 3.4: IMPF SUBJ for *volō* and *colō* in Latin

The present infinitive and imperfect subjunctive are formed by respectively adding the suffixes */-se/* and */-sē/* to the combination of the root and the thematic vowel. Rhotacism (*-s- → -r-*) and assimilation (*-s- → -l-*) obscure the overt forms in Table 3.3 and Table 3.4. The crucial fact that these forms demonstrate is that where *colō* has an *-e-* that intervenes between the root and the suffix, *volō* simply appends the suffix directly to the root. Thus, we do not find forms like a present infinitive **velere* or a imperfect subjunctive 1SG **velerem*, 2SG **velerēs*, etc. Other idiosyncratic properties of the verb, such as its present subjunctive, will be returned to later on. For now, we can set these forms aside as they are largely independent from the distribution of athematicity.

3.2.1.2 *NŌLŌ* AND *MĀLŌ* IN THE CLASSICAL PERIOD

The aforementioned idiosyncratic properties of *volō* ‘I want’ are carried over into the present indicative paradigms of the derivatives *nōlō* ‘I do not want’ and *mālō* ‘I prefer’.⁸

⁸That is, the root allomorphy in the second singular, the athematicity, and the irregular vowel quality of the theme vowel in the first plural.

In addition to these characteristics, the paradigms of *nōlō* and *mālō* demonstrate an alternation between ANALYTIC/SEMI-ANALYTIC forms in the second person and third plural and SYNTHETIC forms elsewhere.

NŌLŌ Beginning with the present indicative of *nōlō*, shown below in Table 3.5, the 2SG, 2PL, 3SG forms are expressed analytically using two words, the negative particle *nōn* ‘not’ and the relevant form of *volō*.

	SINGULAR	PLURAL
1ST	<i>nōlō</i>	<i>nōlumus</i>
2ND	<i>nōn vīs</i>	<i>nōn vultis</i>
3RD	<i>nōn vult</i>	<i>nōlunt</i>

Table 3.5: PRES IND of *nōlō* ‘I not want’ in Latin

These forms contrast with the first person and the third person plural, where contraction occurs. Contracted forms like ***nīs*, ***nōlt*, and ***nōltis* are otherwise unattested; these person and number combinations admit analytic forms only. Analytic forms involving *nōn* and a form of *vol-* are occasionally attested for other forms in the paradigm, but these are notably rare in the Classical period. For most of the cells, an analytic combination is completely unattested, a surprising fact given the high frequency of this verb. It is clear that the synchronic distribution of *nōlō* is characterized by a genuine SYNTHETIC-ANALYTIC alternation, as shown in Table 3.5.

It should be noted that the verb *nōlō* participates in a unique construction, the negative imperative, which is expressed through the form *nōlī* ‘do not!’ with an infinitive.⁹ The verb *volō* itself lacks an imperative entirely, precluding the existence of either analytic or synthetic imperatives.

MĀLŌ The verb *mālō* is historically a contraction from *magis volō* ‘I want more’, i.e. ‘I prefer’, which in Classical Latin was expressed via the contraction of a prefixal element

⁹The second person plural form is *nōlīte*.

mā- with *volō*. The paradigm of *mālō* in the Classical period demonstrates a similar distribution to *nōlō*, albeit with a slightly different type of morphological pattern, seen below in Table 3.6.

	SINGULAR	PLURAL
1ST	<i>mālō</i>	<i>mālūmus</i>
2ND	<i>māvīs</i>	<i>māvultis</i>
3RD	<i>māvult</i>	<i>mālunt</i>

Table 3.6: PRES IND of *mālō* ‘I prefer’ in Latin

The second person (*māvīs* and *māvultis*) and third person singular (*māvult*) show an uncontracted prefixal *mā-*, which alternates with a contracted stem *māl-* seen in the other forms of the present indicative. In the 2SG, 2PL, 3SG, the SEMI-ANALYTIC forms of *mālō* parallel the ANALYTIC *nōlō* in that they are not fully contracted; ***mīs*, ***mālt*, and ***māltis* are unattested. Otherwise, the synchronic distribution of present indicative follows directly from *volō*.

Finally, the present infinitive (Table 3.7) and imperfect subjunctive (Table 3.8) of both of these verbs are fully contracted in the Classical period. They exhibit the same athematic properties as seen in Table 3.3 and Table 3.4 for *volō*.

<i>nōl-le</i>	‘to not want’
<i>māl-le</i>	‘to prefer’

Table 3.7: PRES INF of *nōlō* and *mālō* in Latin

	1SG	2SG	3SG	1PL	2PL	3PL
NŌLŌ	<i>nōllem</i>	<i>nōllēs</i>	<i>nōllet</i>	<i>nōllēmus</i>	<i>nōllētis</i>	<i>nōllent</i>
MĀLŌ	<i>māllem</i>	<i>māllēs</i>	<i>māllet</i>	<i>māllēmus</i>	<i>māllētis</i>	<i>māllent</i>

Table 3.8: IMPF SUBJ of *nōlō* and *mālō* in Latin

Hypothetical forms like a present infinitive ***nōlere/**mālere* and imperfect subjunctive forms like 1SG ***nōlerem/**mālerem*, etc. are unattested.

3.2.2 SYNCHRONIC DISTRIBUTION IN OLD LATIN (PLAUTUS)

The Old Latin period (3rd-2nd century BCE) directly preceded the Classical period and was the first stage after some of the most characteristic Latin sound changes took place (e.g. rhotacism and weakening). The grammar during this period largely resembled that of Classical, albeit with a significant amount of variation. For this reason, Old Latin shows many archaisms and provides much of the evidence used to reconstruct the predecessor of Latin. As we will see, the forms during this period provide insights regarding the distribution of these forms that provide more information about the prehistory than the paradigm of the Classical period by itself.

The primary authors of this era include Plautus, Cato, Livius Andronicus, Naevius, Ennius, and Lucilius (Weiss 2020: 25). For the forms relevant to this discussion, the authors largely concur with the distribution in the Classical period, apart from Plautus, from which key insights can be drawn. As such, this section will focus primarily on Plautus, although the occasional form from some of the different authors will be mentioned.

Titus Maccius Plautus (ca. 254 - 184 BCE) was a Roman playwright whose works can be dated to the end of the third and beginning of the second century BCE. Even though much of his material was adapted from earlier Greek comedies, the plays of Plautus were distinctly Latin, both in terms of style as well as language (Fortson, 2008). The size of the Plautine corpus is surprisingly large given the period during which it was originally written. The twenty or so plays consist of around twenty thousand verses and is comparable in length to the works of Homer (Fortson 2008: 10).¹⁰ For our purposes, the relatively large corpus size suggests that any systematic pattern found in these texts

¹⁰The twenty-first play, *Vidularia*, is quite incomplete and only has around one hundred attested lines (Fortson, 2008).

is informative as to the synchronic and diachronic nature of the forms within.¹¹

This section discusses the unique properties of *volō*, *nōlō*, and *mālō* in the Plautine texts. The present indicative of each verb will be discussed first, before reviewing forms in the other tenses and moods. Special attention will be given to the present and imperfect subjunctive, which show ANALYTIC and SEMI-ANALYTIC forms that are unique to Old Latin.

3.2.2.1 VOLŌ IN OLD LATIN (PLAUTUS)

Beginning with the verb *volō*, the verb *volō* in early Latin texts is largely identical to the classical period. The two paradigms are compared in Table 3.9

	OLD LATIN	CLASSICAL LATIN
1SG	<i>volō</i>	<i>volō</i>
2SG	<i>vīs</i>	<i>vīs</i>
3SG	<i>volt</i>	<i>vult</i>
1PL	<i>volumus</i>	<i>volumus</i>
2PL	<i>voltis</i>	<i>vultis</i>
3PL	<i>volunt</i>	<i>volunt</i>

Table 3.9: PRES IND of *volō* during Old Latin and Classical periods

A small phonological difference affects the second plural and third singular where Old Latin has a consistent *-o-* vowel in the root (*voltis* and *volt*), while Classical Latin has *-u-* (*vultis* and *vult*). The change in the Classical Latin forms can be attributed to a regular change where *-o-* was raised to *-u-* before *-l-* in a closed syllable, cf. *sulcus* ‘furrow’ < **solkos*.¹²

¹¹For more information on the nature of Plautine texts, see Fortson (2008) and the sources within.

¹²For the synchronic distribution of the root allomorphs *vol-* and *vul-* in the Classical period (e.g. in the 1SG vs 3SG), I assume that the raising rule is still synchronically active. Thus, the single underlying root is *vol-*. For this reason, I leave it aside in the synchronic analyses in discussed in Section 3.4.

3.2.2.2 NŌLŌ IN OLD LATIN (PLAUTUS)

The Plautine distribution of the verbs *nŏlŏ* and *mālŏ*, on the other hand, differs significantly from the Classical period. Beginning with the paradigm for *nŏlŏ*, the distribution of all the different variants of *nŏlŏ* is summarized in the table in Table 3.10.

	CONTRACTED		PREFIXED W/ <i>ne-</i>		PERIPHRASTIC <i>nŏn</i>
1SG	<i>nŏlŏ</i>	x80	–		<i>nŏn volŏ</i> x2
2SG	–		<i>nevīs</i>	x10	<i>nŏn vīs</i> x9
3SG	–		<i>nevolt</i>	x5	<i>nŏn volt</i> x5
1PL	<i>nŏlumus</i>	x1	–		–
2PL	–		–		<i>nŏn voltis</i> x1
3PL	<i>nŏlunt</i>	x1	–		–

Table 3.10: PRES IND of *nŏlŏ* in Plautus (Counts)

Comparing the various attested forms, we can see that the Plautine distribution generally confirms the opposition between the two different sets of cells (1SG, 1PL, 3PL vs 2SG, 2PL, 3SG).¹³ The main way in which this verb differs from its classical counterpart is in the presence of SEMI-ANALYTIC variants in the present indicative as opposed to ANALYTIC forms. In this paradigm, forms with a prefixed *ne-* occur alongside the the periphrastic forms *nŏn vīs*, *nŏn volt*, and *nŏn voltis*, specifically in the 2SG and the 3SG: *nevīs* and *nevolt*. Thus, we find both SEMI-ANALYTIC and ANALYTIC forms in the 2SG and 3SG. The exact functional difference between the forms with *ne-* and those with *nŏn* is not immediately obvious. It is worth noting that there is a strong tendency for the forms that are prefixed with *ne-* to occur line-finally (*nevīs* 7/10 instances, *nevolt* 3/5 instances), suggesting that the distribution of these variants was at least in part influenced by metrical factors. For the purposes of the present discussion, I assume that both the forms with *ne-* and *nŏn* were accepted variants, the choice of which was

¹³There are two attestations of *nŏn volŏ* in the plays of Plautus. However, I believe there is good reason to separate these from the other forms. The two 1SG attestations, *Amph.* 371 and *Men.* 1084, exhibit two characteristics that make them unique: i) they are clearly emphatic, and ii) they are tmetic. The exceptionality of this type of example will be discussed further in Section 3.2.5.

sensitive to meter. However, as we will see in Section 3.2.5, there were also other factors involved in distinguishing between the SEMI-ANALYTIC and ANALYTIC types.

A crucial observation is that Plautus still marks a difference between the 1SG, 1PL, 3PL and 2SG, 2PL, 3SG. The ANALYTIC and SEMI-ANALYTIC forms are limited to the latter cells, while the SYNTHETIC forms are always confined to the former. In other words, the alternation present in the Classical period was already there in Plautus, albeit with the added complication of the SEMI-ANALYTIC forms. In this sense, the alternation between the SEMI-ANALYTIC and ANALYTIC variants is independent of the larger contrast between ANALYTIC and SYNTHETIC forms.

The other early authors generally confirm the opposition between ANALYTIC and SYNTHETIC forms, with the 2SG, 2PL and 3SG forming ANALYTIC constructions with the negative particle *nōn*. The authors Caecilius and Naevius employ ANALYTIC forms for the 2SG (*nōn vīs* in Caecil. x1, Naev. x1) and the 3SG (*nōn vult* in Caecil. x1).¹⁴

An unexpected contraction 2PL *nōltis* does appear in Caecil. *Com.* 5; I follow the suggestion in Weiss (2020: 458) that this form is likely a nonce formation.¹⁵ Besides this exceptional attestation, the evidence points to a ANALYTIC or SEMI-ANALYTIC outcome in the second plural. While Plautus only shows an ANALYTIC *nōn voltis*, I assume that the second plural also permitted a SEMI-ANALYTIC form with a prefixed *ne-*, i.e. *ne-voltis**, similar to the third singular and second singular; it is simply unattested in the texts.¹⁶

In all, the forms of *nōlō* in the early period largely confirm the contrast between

¹⁴Caecilius also uses the form *noenu volt* (*Com.* 214), which confirms the existence of the non-contracted form of *nōn* < *noenum*, originally from < **ne-qinōm* ‘not one’.

¹⁵The form appears in the following context: *voltis, empty est; nōltis, nōn empty est* ‘if you want her, she’s been sold; if you don’t she hasn’t’ (Caecil. *Com.* 5). The repetitive context may have led to the form being interpreted as comical or otherwise irregular.

¹⁶This assumption is supported by the fact that second plural forms are generally less frequent by my own counts than the other person-number combinations. This may also have to do with the genre of text.

ANALYTIC and SYNTHETIC forms found in the Classical period, albeit with the addition of SEMI-ANALYTIC forms prepended with *ne-*. The two stages are compared below:

	PLAUTUS	CLASSICAL
1SG	<i>nōlō</i>	<i>nōlō</i>
2SG	<i>nōn vīs, nevīs</i>	<i>nōn vīs</i>
3SG	<i>nōn volt, nevolt</i>	<i>nōn vult</i>
1PL	<i>nōlumus</i>	<i>nōlumus</i>
2PL	<i>nōn voltis, nevoltis*</i>	<i>nōn vultis</i>
3PL	<i>nōlunt</i>	<i>nōlunt</i>

Table 3.11: *nōlō* in Plautus and in Classical Latin

3.2.2.3 MĀLŌ IN OLD LATIN (PLAUTUS)

The Plautine distribution of the present indicative of *mālō* differs only slightly from its classical distribution. While its distribution, summarized in Table 3.12, is not as fleshed out as that of *nōlō* in terms of the number of attestations, several conclusions can still be drawn.

	CONTRACTED		PREFIXED W/ <i>mā-</i>		PERIPHRASTIC <i>magis</i>	
1SG	<i>mālō</i>	x7	<i>māvolō</i>	x8	<i>magis volō</i>	x1
2SG	—		<i>māvīs</i>	x7	<i>magis vīs</i>	x1
3SG	—		—		—	
1PL	—		—		—	
2PL	—		<i>māvoltis</i>	x1	—	
3PL	—		—		—	

Table 3.12: PRES IND of *mālō* in Plautus (Counts)

Although the uncontracted form 1SG *māvolō* appears to contradict the expected contrast between SYNTHETIC and SEMI-ANALYTIC forms found in the Classical period, it should be noted that all seven instances of the form are limited to the end of the line, suggesting that meter once again played a role in conditioning its use. Its SYNTHETIC counterpart *mālō*, on the other hand, does not seem to be limited to any specific position. Given that the use of the latter variant was less restricted, it seems the SYNTHETIC variant was the less marked variant of the two. Crucially, the forms 2SG *māvīs* and 2PL

māvoltis are not restricted to line-final position; only one instance of 2SG *māvīs* out of seven occurs line-finally (*Truc.* 153). The lack of metrical limitation for the second person forms suggests that the SEMI-ANALYTIC forms are the less marked forms for these cells. Taken together with *mālō*, it again seems likely that the SYNTHETIC versus SEMI-ANALYTIC distinction was already present during this period. The final forms in Table 3.12 that require explanation are the forms of *volō* that combine ANALYTICALLY with the adjective *magis*. These forms can likely be explained as especially emphatic and are largely independent of the distribution of contraction.¹⁷

3.2.2.4 FORMS OUTSIDE OF THE PRESENT INDICATIVE FOR *NŌLŌ* AND *MĀLŌ*

From what we have seen so far, the present indicative paradigms of *nōlō* and *mālō* in Old Latin do not contradict the distribution in the Classical period; in fact, the attested distributions largely confirm the opposition between SYNTHETIC forms in the 1SG, 1PL, 3PL and ANALYTIC/SEMI-ANALYTIC forms in the 2SG, 2PL, 3SG. In addition, the differences contribute several important pieces of information. The SEMI-ANALYTIC forms of *nōlō* that have a prefixed *ne-* confirm the presence of a more archaic construction of **ne volō*, which will be expanded on in Section 3.2.5. Essentially, the variants *nevīs* and *nevolt* are older forms that were already limited in their use by the time of Plautus. Furthermore, the first person singular form *māvolō* demonstrates that SEMI-ANALYTIC forms of *mālō* were not limited to the second person or third singular in the Plautine period. However, as with the SEMI-ANALYTIC forms of *nōlō* with *ne-*, these forms were limited and likely conditioned by metrical position.

¹⁷The appearance of *magis* with forms of *volō* does suggest that speakers were still synchronically able to associate the verb *mālō* with the combination of *magis* and the respective form of *volō*, despite the phonological rules governing their contraction no longer being transparent. The fact that *mālle* ‘to prefer’ is a contracted form of *magis velle* is also mentioned explicitly by Cicero several centuries later in *Orator* 154.4.

A more significant difference between the Plautine and Classical forms is that the earlier texts demonstrate variation in terms of contraction beyond the present indicative. To reiterate, systematic contrasts between SYNTHETIC and ANALYTIC forms for *nōlō* and SYNTHETIC and SEMI-ANALYTIC *mālō* respectively is limited to the present indicative.¹⁸ For all other tense/mood combinations, the two verbs have contracted outcomes in the Classical period. In Plautus, however, the other tenses and moods show various examples of unexpected ANALYTIC and SEMI-ANALYTIC forms.

The aberrant non-SYNTHETIC forms in Plautus merit their own explanation as their existence questions whether the alternation in the present indicative in Plautus actually constituted a real alternation. If Plautus allowed both SYNTHETIC and ANALYTIC or SEMI-ANALYTIC forms for all members of the paradigm, any ‘pattern’ found in the present indicative may just be a coincidence, rather than a meaningful alternation. This section discusses these unexpected forms and argues that they can be explained as a side effect of the meter and do not contradict the evidence gleaned from the investigation of the forms of the present indicative in Plautus.

3.2.2.5 PRESENT SUBJUNCTIVE

The present subjunctive especially demonstrates clear variation in terms of contraction. The distribution of forms for this mood for *nōlō* and *mālō*, seen in Table 3.13 and Table 3.14 below, suggests that metrical position influenced whether these forms appeared as SYNTHETIC or ANALYTIC/SEMI-ANALYTIC.

¹⁸There are cases of the combination of *nōn* and forms of *volō* in Classical Latin outside of the second person and third singular present indicative; however, these are few and far between. In the present indicative, there are no examples of ANALYTIC or SEMI-ANALYTIC forms outside of the expected 2SG, 2PL, 3SG.

ANALYTIC	MEDIAL	FINAL	SYNTHETIC	MEDIAL	FINAL
<i>nōn velīm</i>	–	–	<i>nōlim</i>	x5	–
<i>nōn velīs</i>	–	x2	<i>nōlīs</i>	x3	–
<i>nōn velīt</i>	–	x2	<i>nōlit</i>	x1	–
<i>nōn velīmus</i>	–	–	<i>nōlīmus</i>	–	–
<i>nōn velītis</i>	–	–	<i>nōlītis</i>	–	–
<i>nōn velint</i>	–	x2	<i>nōlint</i>	–	–

Table 3.13: SYNTHETIC/ANALYTIC forms in the PRES SUBJ of *nōlō*

SEMI-ANALYTIC	MEDIAL	FINAL	SYNTHETIC	MEDIAL	FINAL
<i>māvelīm</i>	x4	x12	<i>mālim</i>	x15	–
<i>māvelīs</i>	x2	–	<i>mālīs</i>	x2	–
<i>māvelīt</i>	x1	–	<i>mālit</i>	–	–
<i>māvelīmus</i>	–	–	<i>mālīmus</i>	–	–
<i>māvelītis</i>	–	–	<i>mālītis</i>	–	–
<i>māvelint</i>	–	–	<i>mālint</i>	x1	–

Table 3.14: SYNTHETIC/SEMI-ANALYTIC forms in the PRES SUBJ of *mālō*

The SYNTHETIC and ANALYTIC forms for *nōlō* show a remarkable pattern. SYNTHETIC forms only occur in medial position, while ANALYTIC forms only occur in verse-final position, meaning that the two types of forms are in complementary distribution. Thus, the SYNTHETIC form *nōlīs* occurs medially while the ANALYTIC form *nōn velīs* occurs in final position. The forms of *mālō* generally follow a similar pattern. Particularly noteworthy is the minimal pair 1SG *māvelīm* and 1SG *mālim*; the SEMI-ANALYTIC form occurs primarily verse-finally (twelve out of sixteen occurrences) while the SYNTHETIC form only occurs in medial position (all thirteen occurrences).¹⁹

In contrast, the distribution of the present subjunctive of *volō* is more evenly dispersed in terms of verse position, despite a notable preference for verse-finality, seen below:

¹⁹The full counts of all relevant forms as well as their attestations in Plautus are provided in Appendix A.2.

	MEDIAL	VERSE-FINAL
<i>velim</i>	x21	x39
<i>velīs</i>	x16	x24
<i>velit</i>	x10	x11
<i>velīmus</i>	x1	–
<i>velītis</i>	x1	–
<i>velint</i>	x4	x3

Table 3.15: PRES SUBJ of *volō* in Plautus (Counts)

This data suggests that some factor conditioned the position in which the forms would preferably be placed in the line. For *nōlō* and *mālō*, the SYNTHETIC forms strongly disprefer being in final position. The ANALYTIC and SEMI-ANALYTIC forms, on the other hand, prefer being in final position. Crucially, the forms of *volō* pattern mostly with the ANALYTIC and SEMI-ANALYTIC forms of *nōlō* and *mālō* specifically. Before accounting for this fact, however, I present the distribution in the other tenses and moods, as they provide additional information.

3.2.2.6 IMPERFECT SUBJUNCTIVE

The imperfect subjunctive has fewer attestations overall and does not support the pattern seen in the present subjunctive (Table 3.16).

	MEDIAL	FINAL
1SG	<i>nōllem</i> (x1), <i>nōn vellem</i> (x1), <i>māvellem</i> (x8)	<i>māvellem</i> (x1)
2SG	<i>nōn vellēs</i> (x1)	–
3SG	–	–
1PL	–	–
2PL	–	–
3PL	–	–

Table 3.16: IMPF SUBJ of *nōlō* and *mālō* in Plautus (Counts)

As seen above, the verb *nōlō* only has one attested SYNTHETIC form, the 1SG *nōllem* (*Stich.* 513), alongside two ANALYTIC forms, the 1SG *nōn vellem* (*Cist.* 506) and 2SG *nōn*

velles (*Aul.* 286), all of which occur in medial position. As for *mālō*, no SYNTHETIC form is attested, while the SEMI-ANALYTIC 1SG *māvellem* is attested nine times, of which only one occurs at the end of the verse. Any other SYNTHETIC or ANALYTIC/SEMI-ANALYTIC variants are otherwise unattested. The minimal data here suggests that there is no form of the imperfect subjunctive that has a strong preference for being verse-final, regardless of the type.

This fact is also supported by the distribution of the imperfect subjunctive of *volō*, where there are very few instances of a verse-final imperfect subjunctive, as seen below in Table 3.17.

	MEDIAL	FINAL
<i>vellem</i>	x14	x1
<i>vellēs</i>	x6	x1
<i>vellet</i>	x4	–
<i>vellēmus</i>	–	–
<i>vellētis</i>	–	–
<i>vellent</i>	x1	–

Table 3.17: IMPF SUBJ of *volō* in Plautus (Counts)

3.2.2.7 OTHER ADDITIONAL ANALYTIC/SEMI-ANALYTIC FORMS

Otherwise, there are only a few isolated examples of non-SYNTHETIC forms of either *nōlō* or *mālō*, none of which suggest any sort of pattern in and of themselves. There is one isolated occurrence of an ANALYTIC 1SG perfect *nōn voluī* (*Mil.* 1409) in medial position, alongside a SYNTHETIC 1SG *nōluī* (all occurrences of which are verse-final) and 3SG *nōluit* (seven out of eight occurrences of which are verse-final). A 3SG future form *māvolet* (*Asin.* 121) is attested in verse-final position. Finally, the only present infinitive form attested is the SYNTHETIC form *nōlle*, attested three times in medial position; an infinitive for *mālō* is unattested.

3.2.2.8 INTERIM SUMMARY

The general preferences that ANALYTIC, SEMI-ANALYTIC, and SYNTHETIC forms of *nōlō*, *mālō*, and *volō* have in terms of medial or verse-final position are summarized below:

	MEDIAL	FINAL
ANALYTIC PRES SUBJ (e.g. <i>nōn velim</i> , etc.)	✗	✓
SEMI-ANALYTIC PRES SUBJ (e.g. <i>māvelim</i> , etc.)	✗	✓
PRES SUBJ of <i>volō</i> (e.g. <i>velim</i> , etc.)	✗	✓
SYNTHETIC PRES SUBJ of <i>nōlō</i> / <i>mālō</i> (e.g. <i>nōlim</i> , <i>mālim</i> , etc.)	✓	✗
ANALYTIC IMPF SUBJ (e.g. <i>nōn vellem</i> , etc.)	✓	✗
SEMI-ANALYTIC IMPF SUBJ (e.g. <i>māvellem</i> , etc.)	✓	✗
IMPF SUBJ of <i>volō</i> (e.g. <i>vellem</i> , etc.)	✓	✗
SYNTHETIC IMPF SUBJ of <i>nōlō</i> / <i>mālō</i> (e.g. <i>nōllem</i> , <i>māllem</i> , etc.)	(✓)	✗

Table 3.18: Preference for forms of the PRES and IMPF SUBJ in Plautus

Table 3.18 demonstrates that the present subjunctive and imperfect subjunctive are generally found in verse final and medial position respectively. However, the SYNTHETIC present subjunctive of *nōlō* and *mālō* disrupt this picture, indicating that something besides the tense of these forms is at play. The puzzle now is identifying what factor conditions these preferences.

It is clear that the preference for medial or verse-final position is not simply conditioned by whether a form is SYNTHETIC or ANALYTIC. Therefore, we must look elsewhere. In what follows, I posit that the forms that prefer medial position can all be distinguished from those that prefer the final position in terms of the phonological weight of the second to last syllable. In the present subjunctive forms of all verb *besides* the SYNTHETIC forms of *nōlō* and *mālō*, the second to last syllable is generally not ‘heavy’. The next section will expand on what ‘heavy’ means and how these forms relate to the general principles of Plautine meter. In the course of the following discussion, I will provide a novel explanation for the distribution of the aforementioned metrical positions.

3.2.3 VERSE-FINALITY

This section presents a novel account for the distribution of forms of *nōlō*, *mālō*, and *volō* in the present subjunctive and imperfect subjunctive. Crucially, this account also explains the difference between the ANALYTIC and SEMI-ANALYTIC of *nōlō* (3SG *nōn volt* versus 3SG *nevolt*) discussed in Section 3.2.2.2.

The distribution of forms discussed in the previous sections can be attributed directly to the meter used in Plautine plays. The two most common meters in Plautus are trochaic septenarius (x8800) and iambic septenarius (x8000), the basic patterns of which are shown below in (1) (Fortson, 2008).²⁰ There is a lot of variation among the feet that are permitted in these lines; different types of feet may replace others within this schema. However, any variation is consistently neutralized at the end of a verse. Crucially, the last full trochee and iamb respectively for the two meters are not allowed to be substituted by different types of feet. This means that the second to last syllable is effectively always a light syllable.

- (1) a. Trochaic Septenarius template
 — ∪ | — ∪ | — ∪ | — ∪ || — ∪ | — ∪ | — ∪ | × ||
 b. Iambic Septenarius template
 ∪ — | ∪ — | ∪ || — | ∪ — | ∪ — | ∪ × ||

This requirement significantly limits what forms are permitted at the end of the verse. For forms of *nōlō* and *mālō*, this has various consequences. In the present subjunctive, all SYNTHETIC forms (e.g. 1SG *nōlim* and 1SG *mālim*) are highly dispreferred, as the contracted syllable is heavy. However, for the ANALYTIC (e.g. 1SG *nōn velim*) and SEMI-

²⁰For reference, the symbols —, ∪ refer to heavy and light syllables respectively, while × refers to an *anceps*, a position that can be either heavy or light. A heavy syllable either contains a long vowel and/or has a coda, while a light syllable has a short vowel with no coda.

ANALYTIC (e.g. 1SG *māvelim*) forms, the second to last syllable remains light, allowing for the insertion of the form in verse-final position.²¹

For the imperfect subjunctive, the preference is reversed, as seen in Table 3.18. The ANALYTIC and SEMI-ANALYTIC forms (e.g. 1SG *nōn vellem* and 1SG *māvellem*) do not occur in verse-final position. This follows directly from the metrical properties discussed above. As the root syllable in *vel-* is always followed by the consonant /-l-/ in the imperfect subjunctive, it consistently scanned as a heavy syllable, preventing its insertion in verse-final position. As the same is true for the forms regardless of whether they were ANALYTIC/SEMI-ANALYTIC or SYNTHETIC, no clear determination about which form was unmarked can be made.

The insights based of the metrical facts can also be extended to explain some of the distributional facts of the present indicative. For example, of the SYNTHETIC forms of the present indicative of *nōlō* and *mālō*, only the 1PL (*nōlumus* and 1PL *mālumus*) are permitted verse-finally. This follows from the fact that the second-to-last syllable was light. As the initial long syllable of the disyllabic reflexes would fall on the second-to-last syllable, the other SYNTHETIC forms of the present indicative would have been illicit at the end of the line for both the trochaic septenarius and the iambic septenarius meters (e.g. 1SG *nōlō*, *mālō* and 3PL *nōlunt*, *mālunt*).

The most significant way that meter can be applied to the present indicative regards the distinction between the ANALYTIC forms of *nōlō* (e.g. 3SG *nōn volt*) and the SEMI-ANALYTIC forms of *nōlo* (e.g. 3SG *nevolt*). That we do not expect heavy syllables in the second-to-last position makes a specific prediction regarding the distribution of the forms of the 2SG and 3SG which show SEMI-ANALYTIC behavior. Since the syllable *nōn* is heavy, we expect the ANALYTIC forms to be replaced in verse-final position by the

²¹The same holds for the forms of *volō* generally.

forms prefixed with *ne-*, as *ne-* is a light syllable. This distribution appears to be borne out by the textual evidence Table 3.19. As expected, the ANALYTIC forms with *nōn* do not occur finally (given that *nōn* is a heavy syllable), while the SEMI-ANALYTIC forms with *ne-* do (as *ne-* is a light syllable).²²

	MEDIAL	FINAL		MEDIAL	FINAL
<i>nōn vīs</i>	x9	x1	<i>nevīs</i>	x4	x6
<i>nōn volt</i>	x5	x0	<i>nevolt</i>	x2	x3

Table 3.19: 2SG and 3SG singular of *nōlō* in Plautus (Counts)

The observation that ANALYTIC/SEMI-ANALYTIC variants exist outside of the present indicative suggests that historically there was an opposition between ANALYTIC and SYNTHETIC forms throughout the paradigm, not just in the present indicative. However, in the Plautine texts, their distribution was largely conditioned by metrical position. The distribution between SYNTHETIC and ANALYTIC/SEMI-ANALYTIC forms in the present indicative, however, remains unexplained.

3.2.3.1 INTERIM SUMMARY

These distributions, combined with our understanding of Plautine meter, allow us to define the synchronic state of variation attested in the texts. For both *nōlō* and *mālō*, the distribution in Plautus was largely identical to that in the Classical period, apart from the presence of SEMI-ANALYTIC forms prefixed with *ne-* for *nōlō*. These latter forms represent an intermediary archaism and were fully eliminated shortly after the time of Plautus. Metrical position otherwise conditioned the appearance of ANALYTIC and SEMI-ANALYTIC forms of both *nōlō* and *mālō* in forms besides the second person and third person singular present indicative.

²²This would predict that a potential SEMI-ANALYTIC form 2PL *nevoltis** would not be licit verse-finally. This might explain its absence from the texts, although its ANALYTIC counterpart *nōn voltis* is not well-attested either way (only a single occurrence).

Taken together, the paradigms of both verbs in Classical Latin were the end result of a long process of elimination of ANALYTIC and SEMI-ANALYTIC forms. While many of the variants were still present during the early period, their use was limited. For Plautus, their distribution was governed by the meter; however, their very presence in the Plautine texts indicates that ANALYTIC/SEMI-ANALYTIC variants outside of the present indicative originally exceeded the attested distribution.

Overall, the diachronic development between Plautus and the Classical period can be summarized as follows for *nōlō* (2) and *mālō* (3):

(2)	<i>nōlō</i> ‘I do not want’		(3)	<i>mālō</i> ‘I prefer’	
	SINGULAR	PLURAL		SINGULAR	PLURAL
1SG	<i>nōlō</i>	<i>nōlumus</i>	1SG	<i>mālō</i>	<i>mālumus</i>
2SG	<i>nōn vīs,</i>	<i>nōn voltis,</i>	2SG	<i>māvīs</i>	<i>māvoltis</i>
	<i>nevīs</i>	<i>nevoltis*</i>	3SG	<i>māvolt*</i>	<i>mālunt</i>
3SG	<i>nōn volt,</i>	<i>nōlunt</i>			
	<i>nevolt</i>				↓
		↓			
	SINGULAR	PLURAL		SINGULAR	PLURAL
1SG	<i>nōlō</i>	<i>nōlumus</i>	1SG	<i>mālō</i>	<i>mālumus</i>
2SG	<i>nōn vīs</i>	<i>nōn vultis</i>	2SG	<i>māvīs</i>	<i>māvultis</i>
3SG	<i>nōn vult</i>	<i>nōlunt</i>	3SG	<i>māvult</i>	<i>mālunt</i>

3.2.4 SIMILAR IRREGULAR PATTERNS ACROSS LATIN

This section lays out how *volō* (and its derivatives) relate to a set of other irregular verbs in Latin in terms of their morphophonological expression. As discussed in Section 3.2.1, one of the characteristics of *volō* is that its 2SG, 2PL and 3SG forms are athematic. That

is, there is no overt theme vowel intervening between the root and the personal ending denoting the person and number of the subject of the verb. Several other irregular verbs show similar morphophonological patterns that target the exact same paradigmatic distribution. To prefigure the discussion of each of these verbs, the general pattern is as follows:

- (4) a. The verbs *volō* ‘I want’ (and its derivatives *nōlō* and *mālō*), *ferō* ‘I bear’, *sum* ‘I am’ (and its derivative *possum* ‘I am able’), and *edō* ‘I eat’ are athematic in the second person and third person singular present indicative, the infinitive, and imperfect subjunctive.
- b. This pattern is secondarily accompanied by suppletion in *sum* ‘I am’ (and its derivative *possum* ‘I am able’) between the allomorphs *es-* and *su-* and *edō* ‘I eat’ between the allomorphs *ed-* and *ēs-*.

This section will proceed by outlining the pattern in *ferō* ‘I bear’, comparing its irregular athematic properties with the third conjugation verb *gerō* ‘I bear’, which shows the expected thematic distribution. Afterwards, I discuss how this pattern intersects with suppletive root allomorphy in *sum* ‘I am’ and *edō* ‘I eat’.

3.2.4.1 ATHEMATICITY IN *FERŌ*

Apart from several irregularities such as the aforementioned athematicity and root allomorphy in the perfect, the verb *ferō* ‘I bear’ effectively conjugates like a third conjugation verb.²³ Table 3.20 compares the present active indicative of *ferō* with a regular thematic verb *gerō* ‘I bear’ that shows the expected pattern of thematicity. Crucially,

²³For context, while many Latin verbs are characterized by various degrees of idiosyncratic behavior in perfect forms, *ferō* has a fully suppletive stem *tul* that alternates with *fer* in the *infectum*. See Appendix B.1.1.

the expected forms 2SG **feris*, 3SG **ferit*, and 2PL **feritis* are completely unattested throughout the history of Latin.²⁴

	FERŌ	GERŌ
1SG	<i>fer-ō</i>	<i>ger-ō</i>
2SG	<i>fer-s</i>	<i>ger-i-s</i>
3SG	<i>fer-t</i>	<i>ger-i-t</i>
1PL	<i>fer-i-mus</i>	<i>ger-i-mus</i>
2PL	<i>fer-tis</i>	<i>ger-i-tis</i>
3PL	<i>fer-u-nt</i>	<i>ger-u-nt</i>

Table 3.20: PRES IND of *ferō* and *gerō* in Latin

This athematicity is carried over into both the present infinitive (Table 3.21) and the imperfect subjunctive (Table 3.22).²⁵

<i>fer-re</i>	‘to bear’
<i>ger-e-re</i>	‘to bear’

Table 3.21: PRES INF of *ferō* and *gerō* in Latin

	1SG	2SG	3SG	1PL	2PL	3PL
FERŌ	<i>fer-re-m</i>	<i>fer-rē-s</i>	<i>fer-re-t</i>	<i>fer-rē-mus</i>	<i>fer-rē-tis</i>	<i>fer-re-nt</i>
GERŌ	<i>ger-e-re-m</i>	<i>ger-e-rē-s</i>	<i>ger-e-re-t</i>	<i>ger-e-rē-mus</i>	<i>ger-e-rē-tis</i>	<i>ger-e-re-nt</i>

Table 3.22: IMPF SUBJ of *ferō* and *gerō* in Latin

The other forms of the *infectum* of both verbs pattern together. That is, for the imperfect indicative, future indicative, and present subjunctive, *ferō* and *gerō* pattern together as regular third conjugation verbs. The irregularity of this verb in its imper-

²⁴It has been suggested that these athematic forms are the result of thematic forms that have undergone syncope, e.g. *fers* and *fert* from **feresi* and **fereti* respectively (cf. Weiss (2020: 432), Leumann et al. (1977: 530), etc.) However, as the paradigm for *gerō* indicates, such a syncope rule is certainly not synchronically active.

²⁵In pedagogical circles, the imperfect subjunctive is often said to be ‘derived’ from the present infinitive, given the similarity between the ending *-re* (underlyingly */-se/*) in the present infinitive and the affix *-rē* (underlyingly */-sē-/*) in the imperfect subjunctive. What is interesting for our purposes is the fact that the presence of a theme vowel (and the quality of any such thematic vowel) patterns together across the two types of constructions.

fective forms is exclusively limited to the aforementioned cells (2SG, 3SG, 2PL) in the present indicative, the imperfect subjunctive, and the present infinitive, suggesting that this pattern is not a surface-level coincidence. The fact that these cells precisely match the distribution of athematicity in the paradigm of *volō* further cements the remarkable nature of this pattern.

3.2.4.2 IRREGULARS WITH SUPPLETION AND ATHEMATICITY

This cell-specific distribution of athematicity is also present in the paradigms for *sum* ‘I am’ and *edō* ‘I eat’, accompanied by root suppletion. As Table 3.23 shows,²⁶ these verbs are characterized with root suppletion that follows the exact same distribution as the athematic properties, highlighted in blue.²⁷ The distributions of the suppletive roots *su-/es-* and *ed-/ēs-* are not conditioned by any obvious synchronic phonological alternation and must be memorized by learners.

	SUM	EDŌ	POSSUM
1SG	<i>sum</i>	<i>edō</i>	<i>possum</i>
2SG	<i>es</i>	<i>*ēs</i>	<i>potes</i>
3SG	<i>est</i>	<i>ēst</i>	<i>potest</i>
1PL	<i>sumus</i>	<i>edimus</i>	<i>possumus</i>
2PL	<i>estis</i>	<i>ēstis</i>	<i>potestis</i>
3PL	<i>sunt</i>	<i>edunt</i>	<i>possunt</i>

Table 3.23: PRES IND of *sum*, *edō*, and *possum* in Latin

The distribution for *sum* ‘I am’ in Table 3.23 is mirrored by the verb *possum*, a contraction of the adjective *potis* ‘able’ and *sum* ‘I am’. In the present indicative, the verb alternates between a surface allomorph *pos-* (first person and third plural) and

²⁶The form **ēs* is not directly attested, although the form *comēs* ‘you eat up’ (with the preverb *con-*) is attested in Plautus (*Most.* 12).

²⁷The characterization of the forms of *esse* ‘to be’ with the stem *su-* as athematic is not uncontroversial. After all, the ending *-m* in the first singular is notably a reflex of the old athematic ending **-mi* and is the only verb that demonstrates this reflex. However, as we will see, the distribution of forms indicates that the form was synchronically considered ‘thematic’ in the attested period in one way or another.

pot- (second person and third singular). This can synchronically be derived from an underlying stem *pot-* that undergoes assimilation preceding the initial *s-* of the suppletive root *sum*.

The present infinitive (Table 3.24) and imperfect subjunctive (Table 3.25) of these verbs parallel the distribution of *ferō* (and *volō*). The forms are athematic throughout and have the same root allomorph as the highlighted cells in the present indicative (Table 3.23).²⁸

es-se ‘to be’
ēs-se ‘to eat’

Table 3.24: PRES INF of *sum* and *edō* in Latin

	1SG	2SG	3SG	1PL	2PL	3PL
SUM	<i>essem</i>	<i>essēs</i>	<i>esset</i>	<i>essēmus</i>	<i>essētis</i>	<i>essent</i>
EDŌ	<i>ēssem</i>	<i>ēssēs</i>	<i>ēssset</i>	<i>ēssēmus</i>	<i>ēssētis</i>	<i>ēssent</i>

Table 3.25: IMPF SUBJ of *sum* and *edō* in Latin

It should be noted that the alternation of root allomorphy and athematicity in the verb *ēsse* ‘to eat’ is eventually replaced through a thematicization of the alternants (favoring the root allomorph *ed-*) over the course of the Classical Period, e.g. *ēsse* → *edere*, *ēs* → *edis*, *ēst* → *edit*, etc.

This trend of thematicization has led many learners of Latin as well as scholars to assume various thematicized forms for the present and future imperative as well as the present and imperfect subjunctive, specifically for *ferō* and *edō*. However, these thematic forms, if attested at all, only occur late into the Classical Period. The attested record in Plautus and in Classical Latin supports the aforementioned distribution of patterned athematicity. In Plautus, the present indicative passive is confirmed by the forms *fertur* ‘is borne’ and *ēstur* ‘is eaten’, alongside the regular thematicized *feror* ‘I am borne’. The imperfect subjunctive for these verbs is unattested in the passive.²⁹ The

²⁸The length of the vowel in *ēs* is confirmed by Donat. ad Ter. *Andr.* 81 (Weiss 2020: 459).

²⁹Some editions propose *ferrētur* (*Truc.* 798). In the Classical period, these forms are also athematic,

present imperative likewise shows athematic forms in the second person singular/plural: *fer/ferte* ‘bear!’, *ēs/ēste* ‘eat!’. The athematic present passive infinitive *ferri* ‘to be brought’ is also attested (e.g. *Merc.* 752).

For the purposes of this present discussion, the forms for these two verbs can be summarized as follows. Both *ferō* and *edō* attest the same distribution of athematicity as *sum* and *volō* throughout the present indicative, imperfect subjunctive, and present infinitive. I assume that the passive and imperative forms concur with this pattern, despite their relatively poor attestation throughout the Latin corpus.

In all, this section shows that several irregular verbs fit into a general pattern in which several of the cells of the present indicative, demonstrated by the highlighted cells in the paradigm in Table 3.26, the entire imperfect subjunctive, and the present infinitive pattern together in terms of athematicity, and in some cases root allomorphy.³⁰

	SINGULAR	PLURAL
1ST		
2ND		
3RD		

Table 3.26: Patterned athematicity in the PRES IND

3.2.5 THE DIACHRONIC BACKGROUND OF *VOLŌ*, *NŌLŌ*, AND *MĀLŌ*

This section sets the stage for a discussion of the diachronic development of the alternations between SYNTHETIC and ANALYTIC + SEMI-ANALYTIC forms in *nōlō* and *mālō* respectively. I will first outline the relevant diachronic observations concerning *volō* and its derivatives, before discussing the nature of the different SEMI-ANALYTIC, ANALYTIC,

cf. *ēssētur* (Varro *De Lingua Latina* 5.106).

³⁰The source of this alternation is unclear, but it is likely that the fact that the distribution parallels the distribution of the *-o-* alternant of the theme vowel is not a coincidence. It appears as if the forms with *-e-* were reanalyzed as athematic, while the forms with *-o-* were analyzed as thematic. Exactly how this analysis took place is a problem for future work. It should be noted that the reorganization of the present indicative along the lines of partial athematicity was limited to pre-Latin.

and SYNTHETIC variants from a diachronic perspective.

3.2.5.1 DIACHRONIC BACKGROUND OF *VOLŌ*

The diachronic development of the verb *volō* is complex, but for the purposes of the present discussion, we can focus on a small set of attributes.³¹ Given that the present indicative of *volō* is divided between athematic and thematic forms, it has been argued that this verb was originally athematic throughout the present indicative and has subsequently been partially thematicized. The origin of its present tense lies in an athematic Narten present. The presence of athematic forms in Latin and Old Lithuanian (*velmi* ‘wish, want’, *pavelti* ‘allow’) suggest that Proto-Indo-European had an athematic present; however, it is the parallel between the Latin present subjunctive in *vel-ī-* (e.g. 2SG *velīs*), originally from the optative **uel-ih₁-* alongside counterparts in other languages, for example the 2SG form *wileis* ‘want’ in Gothic, that cements the fact that the present of **uel(h₁)-* went back to an athematic present type, specifically a Narten present (Narten, 1968).³² Thus, 2SG *velīs* directly parallels the 2SG *wileis*.³³

In this discussion, I assume that before any contraction took place in the derivatives of this verb, the present indicative had already been divided into two sets of athematic and thematic cells, as seen in the template for patterned athematicity in Table 3.26. Thus, the present tense stem **uel-* was followed by thematic morphology in the 1SG, 1PL, and 3PL, and athematic in the 2SG, 2PL, and 3SG. As we will see, the phonological circumstances that led to contraction in these verbs must have occurred at a much later date than the development of an athematic-thematic alternation.

³¹For a more detailed discussion of the phonological history of this verb, see the discussions in Leumann et al. (1977), Weiss (2020).

³²This present type was characterized by a lengthened grade in the singular active and full grade in the plural, with stress on the root throughout the paradigm.

³³cf. Harðarson (1993), LIV² (2001), Weiss (2020).

Given the idiosyncratic nature of the second singular $v\bar{s}$, its development requires special attention. There are two alternative proposals for the development of this form in the literature. The first is that the 2SG form $v\bar{s}$ is the result of the ‘invasion’ of a suppletive form from the verb $*_{\zeta}ueih_1-$, cf. [Osthoff \(1881\)](#), replacing the root $*_{\zeta}uel(h_1)-$ which supplied the forms in the rest of the paradigm. The resulting form $v\bar{s}$ would then simply follow phonologically from $*_{\zeta}ueih_1-s > *_{\zeta}ueis$. Criticizing the costliness of assuming such a suppletive intrusion, [Cowgill \(1978\)](#) departs from this approach and proposes that the second singular can be traced back to an irregular phonetic development affecting the same stem $*_{\zeta}uel-$ as in the other forms. In this account, the athematic 2SG form (post *i*-apocope) $*_{\zeta}uell < *_{\zeta}uelsi$ was pleonastically reinforced by the regular second singular ending *-s*, $*_{\zeta}uells$, further simplified to $_{\zeta}els$ [weλs].³⁴ In this environment, the sequence *-els-* [eλs] underwent an irregular vocalization and further developed into a diphthong, whence $*_{\zeta}ueis > v\bar{s}$.³⁵ Given that we are concerned with the development of an alternation that occurred after the division of thematic and athematic forms in the present indicative, exactly which account one prefers is immaterial for this discussion. What is crucial is that $v\bar{s}$ continues an athematic form $*_{\zeta}ueis$, either from $*_{\zeta}ueih_1-s$ or $*_{\zeta}uell-s$. This athematic second singular therefore parallels the athematic second singulars found in *ferō* and *edō*, despite its aberrant suppletive shape.

Secondary sound changes further affected the paradigm, the most significant of which was the creation of two allophones of */-l-/:* a palatal [-λ-] before *-i-* and when geminated, and a velar [-ɫ-] found elsewhere (cf. [Leumann et al. 1977](#), [Weiss 2020](#)). As **e* before velar */ɫ/* became *o*, the root $*_{\zeta}uel$ developed several phonologically-conditioned allomorphs depending on the quality of the final lateral. In the present indicative, the

³⁴ An unreinforced variant survived in Latin as the conjunction *vel* ‘or, and, even’ $< *_{\zeta}uell$.

³⁵ At this point, Latin already distinguished between a velar */ɫ/* and a palatal */λ/*, both from **l*. The geminate *-ll-* in the proposed form $*_{\zeta}uell(s)$ would have been palatal, thus distinguishing it from the velar */ɫ/* elsewhere in the paradigm.

vowel became **o*, while the present subjunctive, imperfect subjunctive, and present infinitive retained the characteristic **e* quality of the root. These sound changes thus obscured the e-grade that was present in the entire present indicative, producing an alternation of *vī-* in the second singular and *vol-* elsewhere.

We can summarize the entire assumed development of the pre-Latin paradigm before contraction for *volō* and its derivatives below:

	SINGULAR	PLURAL
1ST	<i>*uolō</i> [uołō] < <i>*uel-ō</i> [uełō]	<i>*uolomos</i> [uołomos] < <i>*uel-o-mos</i> [uełomos]
2ND	<i>*ueis</i> [ueis] < <i>*uei-s</i> [ueis]	<i>*uoltis</i> [uołtis] < <i>*uel-tis</i> [uełtis]
3RD	<i>*uolt</i> [uołt] < <i>*uel-t</i> [uełt]	<i>*uolont</i> [uołont] < <i>*uel-o-nt</i> [uełont]

Table 3.27: Pre-Latin paradigm for *volō*

3.2.5.2 DIACHRONIC BACKGROUND OF *NŌLŌ*

Turning to its derivatives, the ingredients for the contractions affecting *nōlō* and *mālō* are clear, as is the general process of contraction. For *nōlō*, the forms in Table 3.27 were combined with the negative element **ne*, also found in other negated verbs, e.g. *nesciō* ‘I do not know’ < *ne + sciō*.

In cases in which the negative element **ne* and form of **uol-* were unverbated and thus formed a phonological unit, the sequence **neuol-* was subject to various sound changes, leading to contraction. This contraction itself was the result of two changes specifically: i) **e > o / _ ŷ* and ii) **ŷ > ∅ / V_α _ V_α* (between like vowels), e.g. *lavātrīna* (Non. 212.7) ‘place for bathing’ > *lātrīna* (Pl. *Curc.* 580). Following these changes, the development of contraction in *nōlō* can be summarized thusly: **ne + ŷol- > *neuol- > *nouol- > nōl-*.³⁶

³⁶The timing of the change of **e > o / _ ŷ* is not crucial for the earlier stages of this development. In other words, unverbation may have occurred before the change. The loss of *ŷ* between like vowels, however, must have occurred after the vowel changes.

The forms with an underlying $*_{\text{u}}\text{uel-}$ (as opposed to $*_{\text{u}}\text{uol-}$) pose a slight complication. We must either postulate that i) the sequence of $*_{\text{no}}\text{uel-}$ (as in the 1SG present subjunctive *nōlim* ‘I would not want’ versus *velim* ‘I would want’) contracted due to regular sound change or ii) was analogically repaired on the basis of contracted forms in the present indicative (that were themselves the result of regular sound change, e.g. *nōlō* < $*_{\text{no}}\text{uolō}$). The first scenario presupposes that the deletion of $*_{\text{u}}$ did not just occur between two vowels of the exact same quality; rather the quality of $*_{\text{o}}$ and $*_{\text{e}}$ sufficed to trigger the deletion. The exact formulation of this sound change is difficult to determine and is largely irrelevant to our present discussion of *nōlō* as we are not directly concerned with the development of contraction in the forms that are not the present indicative.³⁷ However, as we will see shortly, it will become relevant when discussing the contraction of *magis* ‘more’ and *volō*. An additional discussion on the phonological limitations of contraction can be found in Section 3.3.2.

Finally, in cases where the ANALYTIC constructions remained uncontracted, the negative element $*_{\text{ne}}$ was eventually replaced by *nōn*, itself derived from *noenum* ‘not one’ (Pl. *Aul.* 67) < $*_{\text{ne}}\text{ioinom}$. This can be seen as a stage of renewal in the Jespersen’s Cycle (cf. [Jespersen 1917](#), [Dahl 1979](#)) where the original negative element is reinforced by a secondary element, in this case $*_{\text{oinom}}$ ‘one’. This reinforcement of the negator,

³⁷ Another possible outcome of the sequences $*_{\text{o}}\text{uo-}$ and $*_{\text{o}}\text{ue-}$ in Latin is $-\bar{u}-$ by means of a syncope rule which deleted the the second vowel when followed by a consonant (see [Weiss \(2020: 134\)](#) for a full description of the rule with examples). This rule is responsible for the long $-\bar{u}-$ in cases like $*_{\text{i}}\text{ouos}$ ‘law’ > *iūs* and $*_{\text{i}}\text{ouestos}$ > *iūstus* ‘legal’. The attested forms of *nōlō* clearly did not undergo this change, as we do not find forms of *nōlō* with the sequence $*_{\text{n}}\text{ul-}$. We also do not find the intermediate stage $*_{\text{noul-}}$ in any early attestation where we might expect such a form. The forms of *nōlō* thus inform our understanding of the limits of this sound change in what follows.

I offer two reasons as to why this sound change may not have affected these forms. The first is that the vowel in the root may have been blocked from deletion due to the fact that the environment contains a morpheme boundary ($*_{\text{no}}\text{uol-}$). The second reason is that there may have been some restriction blocking the deletion in the root involving the identity of the root. That is, an alternation of $*_{\text{u}}\text{ol-}/*_{\text{ul-}}$ was blocked because one of the alternants would lack a root vowel. An analogical approach might also be considered, although in such a scenario forms of *nōlō* would have to be restored based on the short vowel in *volō*, rather than another form of *nōlō* directly.

emphasized the difference between the ANALYTIC and SEMI-ANALYTIC forms of *nōlō*. Where the ANALYTIC forms (e.g. 3SG **ne volt*) used to resemble the SEMI-ANALYTIC forms (e.g. 3SG **nevolt*), the difference is now more pronounced (e.g. *nōn volt* versus *nevolt* in Plautus).

3.2.5.3 DIACHRONIC BACKGROUND OF *MĀLŌ*

The univerbation of the adjective *magis* ‘more’ and *volō* generally follows the same pattern of contraction caused by the loss of *-*u-*. When **magis* and **uel-* were univerbated, the form was syncopated to **magzuol-*. This */*gzu-/* cluster was simplified with compensatory lengthening of the preceding vowel.³⁸

At this point, the development of the newly-formed sequence of *māuol-* contracted, parallel to the sequence **nouol-*. The development was presumably also due to the rule which conditioned deletion between like vowels, despite the quality of the vowels differing. There is no obvious analogical source for the contraction, unlike the forms for *nōlō* involving a sequence of **noue-*. As all SEMI-ANALYTIC forms of *mālō* at this stage would have either had **māuol-* as in the present indicative (e.g. 1SG **māuolō*) or **māuel-* as in the present subjunctive (e.g. 1SG **māuelīm*), there was no source for the SYNTHETIC outcome. The lack of analogical pathway implies that the final SYNTHETIC outcome was ultimately due to sound change.

The present subjunctive, imperfect subjunctive, and present indicative of *mālō* face the same challenge as *nōlō*, in the sense that contraction of sequences like **-oue-* and **-āve-* (as in 1SG IMPF SUBJ *māvellem*) is unexpected. If the deletion rule was based on both vowels being identical, or at least similar in terms of backness, the aforementioned sequences would not have undergone deletion. In both verbs, an analogical repair in

³⁸This development is perhaps paralleled by **aps#uellō* ‘I tear away’ > **abzuellō* > *āvellō* and **seks#uiroi* ‘board of six men’ > **segzuīroi* > *sēvirī* (Weiss 2020: 192).

the present subjunctive and imperfect subjunctive is possible, as long as the outcomes in the present indicative followed the described course. As we will see in the following section, I argue that the contraction in these forms did not follow directly from regular sound change; rather they are indeed the result of analogy with the present indicative.

In all, the SYNTHETIC forms in the present indicative of *nōlō* and *mālō* are ultimately the reflexes of regular sound changes governing the unverbated sequence of **ne + uol-* and **magis + uol-* respectively outlined above. These developments taken together allow us to take away one key fact from this development. Whatever the conditioning of the contraction was, the lack of contraction in the second singular of both verbs (*nōn vīs/nevīs* and *māvīs*) was phonologically expected, given that the combination of these vowel sequences do not otherwise cause loss of the *-u-*. In this sense, the alternation between the SYNTHETIC forms in the 1SG (*nōlō* and *mālō*) and the SEMI-ANALYTIC forms in the 2SG (*nevīs* and *māvīs*) follows directly from regular developments and do not need a specific explanation phonological or morphological explanation. However, this is not the case for the third singular and second plural, as the fact that the lack of synthesis in these forms is not obviously reducible to phonological developments.

3.2.5.4 FIXING THE CONTRACTION RULE

The rule of ‘contraction between like vowels’ should therefore be re-examined. One possible way to account for the phonological changes is to assume that the rule depended not on the exact quality of the vowel, but rather a specific property of the vowel, e.g. height or backness. In phonetic terms, this may be described in terms of tongue position. As long as both vowels shared some property (or properties) affecting tongue position before and after the pronunciation of the */-w-/*, the movement of the tongue during the */-w-/* may have been reduced, especially in rapid speech.

For our purposes, the semi-glide /-w-/ may be thought of as similar to the vowel /-u-/ albeit with more narrow constriction in the vocal tract (cf. Ladefoged & Maddieson 1996: 323). The semi-glide therefore has fairly similar formants as its corresponding vowel. The quality of its formants has several consequences regarding the formant transitions in sequences of -V₁-w-V₂-. F1 remains fairly stable, while the higher formants (F2, F3) may undergo significant change depending on the quality of V₁ and V₂ (Lisker, 1957). The changes in the higher formants have therefore been identified as particularly important in identifying the semi-glide /-w-/ (Ainsworth & Paliwal, 1984, *a.o.*).

Given that we are trying to explain cases in which the /-w-/ was deleted, we might expect that deletion of /-w-/ would be linked to smaller formant transitions. If the formant transitions were smaller, then perceptually the sequence could be analyzed as /-V₁V₂-/, eventually leading to contraction. The movement of the tongue would be minimized in sequences of similar vowels, resulting in a decrease in formant transitions, allowing a change based on perception. However, the scope of the sound change in Latin suggests that size of the transitions did not factor in to the deletion of the /-w-/, despite the intuition that the small formant transitions would make learners more likely to delete /-w-/ between vowels. For example, a sequence of /-iwi-/ has a large change in F2 (Lisker 1957: 263), especially compared to sequences with back vowels (e.g. /-awa-/). However, there are clear examples of sequences of /-iwi-/ undergoing deletion in Latin, e.g. *oblīvīscar* ‘I will forget’ > *oblīscar* (Acc. *Trag.* 448). That like vowels always contract seems to put into question the intuition that the deletion was due to the failure to perceive the /-w-/ in terms of formant transitions. In sum, there are two facts that we need to account for: i) /-w-/ was deleted between like vowels, including those which would have had large transitions in the second formant, and ii) /-w-/ was also deleted in cases where the vowels were similar but not identical, specifically in the

sequence /-awo-/.

I propose that this sound change was due to the articulatory undershoot (Bauer 2008: 611) of the tongue in these sequences, facilitated primarily by sequences where the backness/frontness (expressed phonetically in terms of the second formant) of the vowels are similar.³⁹ It is clear from the contraction in cases like /-iwi-/ that the crucial factor determining contraction was that the movements of the tongue before and after the articulation were identical (or very similar) to one another. Given that there is not much of a change in the first formant across the board in sequences of /-V₁wV₂-/, we can infer that the tongue position does not change much on the vertical axis during the pronunciation of the /-w-/. Thus, the primary factor that caused undershoot was the backness/frontness of the two vowels. That is, the undershoot of the tongue was specifically in reference to tongue retraction and not tongue raising. Effectively, the tongue had nothing to ‘undershoot’ in terms of tongue height.

In between two identical high vowels, like /-iwi-/, the significant horizontal movements before and after the /-w-/ motivated the reduction of movement. Throughout the whole sequence, the height of the tongue remained relatively unchanged. For sequences of low vowels, e.g. /-awa-/, the movement of the tongue was larger, but was still relatively low.⁴⁰ It is important to note that the F1 of /-w-/ is lower than that of its vocalic counterpart /-u-/ (Maddieson & Emmorey 1985: 171), further supporting the fact that the changes in F1 were not as crucial as the changes in F2 for motivating the undershoot.

For sequences involving non-like vowels, like -a- and -o-, the tongue height during the transitions was not sufficiently different so as to block the deletion. What is crucial

³⁹Compare the deletion of /-w-/ in Polish between vowels which cause the tongue to undergo a lot of movement in Jaskulala (2014: 64-67).

⁴⁰It should be noted that Jaworski (2021), citing research in Jaworski (2009), Jaworski & Gillian (2011), points out that reduction in intervocalic contexts is amplified between low vowels.

is that both vowels would have similar transitions in terms of retraction of the tongue, motivating the deletion of /-w-/ in these sequences. The tongue position would have been similar before and after the pronunciation of /-w-/ in terms of how far back the tongue was. That is, the difference between the transition to the /-w-/ in the first part of the sequence /-awo-/ and the transition from the /-w-/ to the /-o-/ in the latter half was sufficiently similar due to the relatively small difference in backness. This change can be compared with cases of lenition as discussed in [Bauer \(2008\)](#).

This proposal predicts the following restriction on the sound change. If the transitions before and after the pronunciation of /-w-/ were not identical or at least similar, then there would be no deletion. Such a restriction would apply in two ways: i) vowels that have a significant difference in height from one another should not contract, and ii) vowels that have a significant difference in backness/frontness should also not contract. As for the first outcome, sequences of *-avu-/avi-* do not contract. Likewise, large differences in terms of backness/frontness also preclude contraction, e.g. *noctuvigila* ‘night-watching’. This account therefore predicts that sequences of /-owe-/ and /-awe-/ did not contract due to regular sound change. In this sense, an analogical solution to the subjunctive forms and the present infinitive of both *nōlō* and *mālō* is preferable.

It should also be noted that the process of deletion of /-w-/ along with contraction seemed to have been at least recoverable from a synchronic perspective at the time of the early texts (e.g. Plautus), as the variation present in forms like the first singular *mālō* and *māvolō* indicates (e.g. 1SG PRES SUBJ *māvelim* vs *mālim*). As discussed in Section 3.2.3, in some of these SEMI-ANALYTIC and ANALYTIC variants, the deletion and contraction was undone if required by the meter.

3.2.5.5 ON VARIANTS

The synchronic distribution of SYNTHETIC and SEMI-ANALYTIC/ANALYTIC variants, as laid out in Section 3.2.2, requires some discussion in terms of its implications for the diachronic development of these verbs. The reason for this is twofold. Firstly, even if the selection of variants in Old Latin can be attributed to metrical position, meter in and of itself does not explain the presence of these variants. While it might be possible that Plautus (in the capacity of being a playwright) simply created these forms artificially to serve his metrical needs, it would be preferable to find a diachronic explanation that accounts for these forms holistically. Secondly, there are factors besides meter that explain parts of the distribution, hinting at a more complicated prehistory than has previously been considered. This section argues that the evidence in Plautus points to the presence of two sets of variants, i) a unverbated set (SYNTHETIC) that was used in unemphatic contexts, and ii) an ANALYTIC set that was used in emphatic and certain syntactic contexts. By the time of Plautus, this alternation began to break down, but both sets were available for whatever the meter required.

We may begin with the observation that the Plautine forms *nevīs* and *nevolt* confirm the prehistoric univerbation of *ne* and the verb *volō*. This parallels other verbs in which a negative particle has been unverbated with the verb: *nesciō* ‘I do not know’ < *ne* + *sciō* ‘I know’ and *nequeō* ‘I am unable’ < *neque* ‘and not’ + *eō* ‘I go’. All together, these verbs are subject to separation from the negative particle under certain emphatic and syntactic conditions. That emphasis conditions periphrasis in these verbs is evidenced by cases in which an emphatic particle (5) or an overt pronoun or determiner (6) intercedes between the negative particle and the verb.⁴¹

⁴¹Langslow (2009: 717) cites Pl. *Aul.* 67-9 as an example of tmesis with *nequeō*, with *noenum* ‘not one’ as the negative element: *noenum...queo comminisci* ‘I cannot imagine’.

- (5) a. *nōn ēdepol volō profectō.*
 NEG by.Pollux want.1SG indeed
 ‘By god, I certainly do not want [it]!’ (Pl. *Amph.* 371)
- b. *nōn ēdepol sciō.*
 NEG by.Pollux know.1SG
 ‘By god, I do not know!’ (Pl. *Asin.* 299)
- (6) a. *nōn ambōs volō.*
 NEG both.ACC want.1SG
 ‘I do not want both [of you]!’ (Pl. *Men.* 1084)
- b. *nōn tu scīs?*
 NEG 2SG.NOM know.2SG
 ‘Do you not know?’ (Pl. *Amph.* 703)
- c. *nōn id voluī dicere.*
 NEG 3SG.NEUT.ACC want.PERF.1SG say.INF
 ‘I did not want to say it.’ (Pl. *Cas.* 366)

In addition to cases involving an intervening element, these negative verbs may be separated from the negative element *neque* in coordinated constructions.⁴²

- (7) a. *neque volō neque postulō neque censeō*
 NEG.CONJ want.1SG NEG.CONJ request.1SG NEG.CONJ think.1SG
 ‘I do not want, or request, or think [it]’ (Pl. *Trin.* 696)
- b. *neque sciō neque nōvī*
 NEG.CONJ know.1SG NEG.CONJ know.PERF.1SG
 ‘I do not know or recognize’ (Pl. *Epid.* 576)

These examples demonstrate that, even in Plautus, emphasis and coordination served as conditioning environments of separation of the negative element and the verb, creating ANALYTIC forms.⁴³

⁴²Plautus even allows the combination of *neque* with forms of *queō* ‘I am able’, which itself is a backformation from *nequeō* ‘I am unable’: *neque quīvī* ‘and I was unable’ (Pl. *Rud.* 91).

⁴³This is not to say that tmesis was a requirement of emphatic uses of these verbs. Rather, the emphasis allowed for tmesis in the appropriate contexts. The replacement of *ne* with *nōn* even in non-tmetic environments (i.e. *nōn vīs*, *nōn volt*, etc. confirms the decompositionality in terms of prosodic

From a diachronic perspective, the variation between SYNTHETIC and ANALYTIC forms had several consequences. In unemphatic and most syntactic contexts, the negative element *ne* was unverbated with *volō*, allowing for the newly formed word to undergo sound changes. These SYNTHETIC constructions eventually became the SYNTHETIC forms that we see in most of the cells in the paradigm. In cases of emphasis and certain syntactic conditions like coordinated structures, the verb did not undergo unverbation, allowing for the negative element to be replaced by *nōn*.

Therefore, the predecessor of the SYNTHETIC form 1SG *nōlō* is a unverbated form, while the predecessor of the ANALYTIC form 3SG *nōn vult* ‘s/he does not want’ was a non-unverbated form, from an emphatic or other syntactic context which did prevented unverbation. What is left unexplained so far is exactly how these variants were combined into one singular paradigm.

3.3 DIACHRONIC EXPLANATION

The previous discussion focused on the general synchronic distribution of the relevant forms of *volō*, *nōlō*, and *mālō* as well as the relevant facts regarding their diachronic history. However, it is still unclear exactly why the forms are distributed in the way they are in terms of the SYNTHETIC and ANALYTIC/SEMI-ANALYTIC alternation. The following sections set the stage for determining the diachronic development of these verbal forms. Section 3.3.1 summarizes the general proposed historical development of *nolō* ‘I do not want’. Section 3.3.2 and Section 3.3.3 discuss a phonological and an analogical approach respectively. This section concludes with a brief aside on why I believe usage and frequency are irrelevant to the development of periphrasis in *nōlō* and *mālō*.

parsing, as otherwise one would expect the *ne* to remain as such.

3.3.1 SETTING THE STAGE

This section summarizes the general diachronic development of $n\bar{o}l\bar{o}$, before outlining the main issue of interest, namely the origin of the distribution of SYNTHETIC and ANALYTIC variants within the present indicative. While this section focuses primarily on the development of $n\bar{o}l\bar{o}$, the development of $m\bar{a}l\bar{o}$ can be dealt with *mutatis mutandis*, as the underlying principles are the same.

The starting point of the paradigm of $n\bar{o}l\bar{o}$ (Table 3.28) was ANALYTIC, with a lack of univertation between the negative element and the verb.

	SINGULAR	PLURAL
1ST	$*ne \underset{\sim}{u}ol\bar{o}$	$*ne \underset{\sim}{u}olomos$
2ND	$*ne \underset{\sim}{u}eis$	$*ne \underset{\sim}{u}oltis$
3RD	$*ne \underset{\sim}{u}olt$	$*ne \underset{\sim}{u}olont$

Table 3.28: Pre-univertated paradigm for $*ne \underset{\sim}{u}ol\bar{o}$

The verbal part of the construction was identical to the respective form of $\underset{\sim}{u}ol\bar{o}$. I assume that, at this stage, the verb had already acquired the distribution of athematic and thematic forms, as well as the special stem in the second singular. Accordingly, the ANALYTIC ancestor of $n\bar{o}l\bar{o}$ also had this distribution. It should be noted that there was nothing special about the combination of $*ne$ and $\underset{\sim}{u}ol\bar{o}$ itself when compared to any other verb in a negated context.

STAGE 1: The first change from this stage, seen in (8), involved the development of two separate variants, one of which was univertated while the other remained non-univertated, identical with the original paradigm in Table 3.28.

(8) Stage 1: Development of Univerbation:

			Non-univerbated	
			SINGULAR	PLURAL
			1ST	<i>*ne uolō</i> <i>*ne uolomos</i>
			2ND	<i>*ne ueis</i> <i>*ne uoltis</i>
			3RD	<i>*ne uolt</i> <i>*ne uolont</i>
Pre-univerbated			Univerbated	
SINGULAR	PLURAL		SINGULAR	PLURAL
1ST	<i>*ne uolō</i> <i>*ne uolomos</i>	$\Rightarrow$	1ST	<i>*neuolō</i> <i>*neuolomos</i>
2ND	<i>*ne ueis</i> <i>*ne uoltis</i>		2ND	<i>*neueis</i> <i>*neuoltis</i>
3RD	<i>*ne uolt</i> <i>*ne uolont</i>		3RD	<i>*neuolt</i> <i>*neuolont</i>

At this point, we may speak of an alternation between SEMI-ANALYTIC forms (the univerbated combination) and ANALYTIC forms. The nature of this alternation is as discussed in Section 3.2.5.5. The SEMI-ANALYTIC complex of the negative particle and the verb was used in contexts which were unemphatic, which presumably was most of the time the verb was used. The sequence, however, was not univerbated in emphatic contexts, as well as certain syntactic contexts, such as cases involving negative conjunction (cf. Section 3.2.5.5). The fact that Latin inherited two paradigms is confirmed by the presence of both non-univerbated and univerbated alternants in Plautus: *nōn volō* < **ne uolō* and *nōlō* < **neuolō*.

STAGE 2: The SEMI-ANALYTIC variant, as it formed one prosodic unit, was subject to various sound changes, the first of which was the change of **e* > *o* preceding **-u-*. This is demonstrated in (9), where the SEMI-ANALYTIC variant underwent a vowel change in the first syllable.

(9) Stage 2: $*e > o / \text{---} \underset{\sim}{u}$:

Non-Univerbated				Non-Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*ne_uolō</i>	<i>*ne_uolomos</i>	⇒	1ST	<i>*ne_uolō</i>	<i>*ne_uolomos</i>
2ND	<i>*ne_ueis</i>	<i>*ne_uoltis</i>		2ND	<i>*ne_ueis</i>	<i>*ne_uoltis</i>
3RD	<i>*ne_uolt</i>	<i>*ne_uolont</i>		3RD	<i>*ne_uolt</i>	<i>*ne_uolont</i>

Univerbated				Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*ne_uolō</i>	<i>*ne_uolomos</i>	⇒	1ST	<i>*no_uolō</i>	<i>*no_uolomos</i>
2ND	<i>*ne_ueis</i>	<i>*ne_uoltis</i>		2ND	<i>*no_ueis</i>	<i>*no_uoltis</i>
3RD	<i>*ne_uolt</i>	<i>*ne_uolont</i>		3RD	<i>*no_uolt</i>	<i>*no_uolont</i>

Here, the non-univerbated variant remained unchanged. This is the case throughout most of the course of the development, as it is only in the later stages that the ANALYTIC set undergoes changes.

STAGE 3: At this point in the development of these forms, the first person forms and the third person plural were differentiated from the other forms as they underwent contraction in the first and second syllable, due to loss of $*-\underset{\sim}{u}-$ between the two like vowels; unexpectedly this sound change did not seem to affect the 3SG and 3PL. This change is represented in (10).

(10) Stage 3: Contraction following $\underset{\sim}{u}$ -loss between like vowels:

Non-Univerbated				Non-Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*ne</i> $\underset{\sim}{u}$ olō	<i>*ne</i> $\underset{\sim}{u}$ olomos	$\Rightarrow$	1ST	<i>*ne</i> $\underset{\sim}{u}$ olō	<i>*ne</i> $\underset{\sim}{u}$ olomos
2ND	<i>*ne</i> $\underset{\sim}{u}$ eis	<i>*ne</i> $\underset{\sim}{u}$ oltis		2ND	<i>*ne</i> $\underset{\sim}{u}$ eis	<i>*ne</i> $\underset{\sim}{u}$ oltis
3RD	<i>*ne</i> $\underset{\sim}{u}$ olt	<i>*ne</i> $\underset{\sim}{u}$ olont		3RD	<i>*ne</i> $\underset{\sim}{u}$ olt	<i>*ne</i> $\underset{\sim}{u}$ olont
Univerbated				Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*nou</i> $\underset{\sim}{u}$ olō	<i>*nou</i> $\underset{\sim}{u}$ olomos	$\Rightarrow$	1ST	<i>*nō</i> lō	<i>*nō</i> lomos
2ND	<i>*nou</i> $\underset{\sim}{u}$ eis	<i>*nou</i> $\underset{\sim}{u}$ oltis		2ND	<i>*nou</i> $\underset{\sim}{u}$ eis	<i>*nou</i> $\underset{\sim}{u}$ oltis
3RD	<i>*nou</i> $\underset{\sim}{u}$ olt	<i>*nou</i> $\underset{\sim}{u}$ olont		3RD	<i>*nou</i> $\underset{\sim}{u}$ olt	<i>*nō</i> lont

Crucially, the other SEMI-ANALYTIC forms (2SG, 3SG, and 2PL) in the paradigm did not undergo contraction. The second singular remained SEMI-ANALYTIC due to the fact the *-o-* and *-ei-* would not contract; the loss of $\underset{\sim}{u}$ - simply did not occur, as the conditioning environment was not correct. The circumstances of why the third singular and second plural did not contract are not immediately obvious, although several potential explanations come to mind. One may note that the second plural and third singular contain heavy syllables, while the first person forms and the third person plural have open syllables. This fact might imply a phonological factor. Additionally, the forms which did not contract fall conspicuously in the same distribution as the irregular distribution exhibited by the verbs discussed in Section 3.2.4, implying that analogy may have played a role.

As we will see shortly, the differentiation between SEMI-ANALYTIC and SYNTHETIC forms in the univerbated paradigm at this stage ultimately determined the distribution

we see in Plautus and in the Classical Period. In other words, whatever prevented the second plural and third singular from undergoing contraction is responsible for the SYNTHETIC-ANALYTIC alternation. Thus, identifying why this differentiation occurred is crucial in understanding the history of these verbs.

INTERIM SUMMARY: Taking stock of what we have discussed right now, the verb *nōlō* at this stage engages in two different types of alternations. The first is an alternation between ANALYTIC forms in emphatic contexts and SEMI-ANALYTIC/SYNTHETIC elsewhere. The other alternation takes place within the latter paradigm and consists of SEMI-ANALYTIC and SYNTHETIC forms. This situation contrasts with the attested pattern in Old and Classical Latin where the SEMI-ANALYTIC forms have been replaced with ANALYTIC forms.

It should be noted that if we compare the paradigm of unverbated forms of *nōlō* in Stage 3 with that of *mālō* during the Old and Classical period, we can see that the type of alternation is the same, namely between SYNTHETIC forms in the 1SG, 1PL, and 3PL and SEMI-ANALYTIC forms in the 2SG, 2PL, and 3SG. In this way, the origin of the alternations in both *nōlō* and *mālō* are the same. As we will see shortly, the fact that the attested Latin alternations are not the same between the two verbs is because the replacement of SEMI-ANALYTIC forms with ANALYTIC forms in *nōlō* has obscured the original distribution.

STAGE 4: Moving on to the next stage, seen in (11), the sequence **no-* underwent a simple analogical renewal with *ne-* in the SEMI-ANALYTIC forms that did not originally contract.

(11) Stage 4: Analogical restoration of the negative element *ne*:

Non-Univerbated				Non-Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*ne</i> <i>uolō</i>	<i>*ne</i> <i>uolomos</i>	⇒	1ST	<i>*ne</i> <i>uolō</i>	<i>*ne</i> <i>uolomos</i>
2ND	<i>*ne</i> <i>ueis</i>	<i>*ne</i> <i>uoltis</i>		2ND	<i>*ne</i> <i>ueis</i>	<i>*ne</i> <i>uoltis</i>
3RD	<i>*ne</i> <i>uolt</i>	<i>*ne</i> <i>uolont</i>		3RD	<i>*ne</i> <i>uolt</i>	<i>*ne</i> <i>uolont</i>
Univerbated				Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	<i>*nōlō</i>	<i>*nōlomos</i>	⇒	1ST	<i>*nōlō</i>	<i>*nōlomos</i>
2ND	<i>*noue</i> <i>is</i>	<i>*nouolt</i> <i>is</i>		2ND	<i>*neue</i> <i>is</i>	<i>*neuolt</i> <i>is</i>
3RD	<i>*nouolt</i>	<i>*nōlont</i>		3RD	<i>*neuolt</i>	<i>*nōlont</i>

This assumption can be motivated by the fact that the sequence **no-* would not have otherwise marked negation in the language, as the change from **e* to **o* did not affect the negative particle elsewhere. For example, other univerbated negative verbs, such as *nequeō* ‘I am unable’, would have had the regular form *ne-* prefixed to the verb at this stage, rather than **no-*. The analogical replacement is therefore the simple application of the regular form **ne*.

STAGE 5: After this replacement, the diphthong **-ei-* ultimately monophthongized, resulting in a second singular *vīs* (*Merc.* 389) and *nevīs* (*Curc.* 82).⁴⁴ After this point,

⁴⁴I have abstracted away from some vocalic developments involving weakening, as these sound changes are irrelevant to the development of the two variants. This includes the change to *-u-* of the theme vowel in the first plural forms: *volumus* ‘we want’ instead of **volimus*.

With specific regard to the monophthongization of *-ei-*, see Weiss 2020: 110. Plautus still had *ei-*, as shown by *eiram* (*Truc.* 264). Evidently, however, the monophthongization had already occurred in *volō* and its derivatives, as **veis* is unattested.

There remains the question why the sequence **neueis* did not contract after the restoration of **ne*. One possibility is that the diphthong in the second syllable prevented the contraction, either due to not

the general negative marker **ne* was replaced by the form *nōn* < *noenum* ‘not one’, affecting the ANALYTIC forms across the board, shown below in (12).

(12) Stage 5: Analogical replacement of the negative element *nōn*:

Non-Univerbated			Non-Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL
1ST	<i>*ne</i> <i>uolō</i>	<i>*ne</i> <i>uolomos</i>	⇒	1ST	<i>*nōn</i> <i>uolō</i> <i>*nōn</i> <i>uolumus</i>
2ND	<i>*ne</i> <i>ueis</i>	<i>*ne</i> <i>uoltis</i>		2ND	<i>*nōn</i> <i>ueis</i> <i>*nōn</i> <i>uoltis</i>
3RD	<i>*ne</i> <i>uolt</i>	<i>*ne</i> <i>uolont</i>		3RD	<i>*nōn</i> <i>uolt</i> <i>*nōn</i> <i>uolunt</i>

Univerbated			Univerbated		
	SINGULAR	PLURAL		SINGULAR	PLURAL
1ST	<i>*nōlō</i>	<i>*nōlomos</i>	⇒	1ST	<i>*nōlō</i> <i>*nōlumus</i>
2ND	<i>*neueis</i>	<i>*neuoltis</i>		2ND	<i>*neueis</i> <i>*neuoltis</i>
3RD	<i>*neuolt</i>	<i>*nōlont</i>		3RD	<i>*neuolt</i> <i>*nōlunt</i>

Eventually the SEMI-ANALYTIC forms with *ne* in the univerbated paradigm were also affected by this analogical restoration. However, unlike the ANALYTIC forms, this process was gradual and may have been restricted to certain sociolinguistic environments. The gradual nature of this change explains the fact that the prefixal *ne-* in other verbs univerbated with *ne-* is variously replaced by *nōn* throughout the early period and into the Classical period, e.g. *non quitast* ‘she couldn’t’ (Terence *Hec.* 572), instead of **nequitast*. Cicero also occasionally has forms of *non queō* instead of *nequeō* (Langslow, 2009, 717).⁴⁵

being similar enough in quality or due to a restriction on syllable weight (cf. Section 3.3.2).

⁴⁵Despite the presence of these examples, it is undeniable that in the Classical period we find *nōn* with *nōlō* in the periphrastic forms across the board. This poses an interesting question regarding the other negated verbs which are univerbated. If *ne-* was replaced in *nōlō*, why not in *nesciō*? A possible solution would be to invoke the fact that only *nōlō* had an alternation between *ne-* with *n-*. This

The gradual replacement of **ne* with *nōn* in the SEMI-ANALYTIC forms is still represented by the distribution of forms found in Old Latin. In the 2SG/2PL forms and 3SG, the forms with **ne* have been replaced with the forms with *nōn*, implying that the distinction between the two paradigms, the emphatic one and the unemphatic one, had already begun breaking down in these forms. While Plautus generally uses the forms with *nōn*, it is clear that the forms with *ne-* were still available, albeit in a limited sense as they are generally conditioned by metrical requirements.⁴⁶ These developments are summarized in (13) below.

(13) Emphatic and Unemphatic Forms in Plautus:

Emphatic			Unemphatic		
	SINGULAR	PLURAL		SINGULAR	PLURAL
1ST	<i>nōn volō</i>	<i>nōn volumus</i>	1ST	<i>nōlō</i>	<i>nōlumus</i>
2ND	<i>nōn vīs</i>	<i>nōn voltis</i>	2ND	<i>nōn vīs/nevīs</i>	<i>nōn voltis/nevoltis*</i>
3RD	<i>nōn volt</i>	<i>nōn volunt</i>	3RD	<i>nōn volt/nevolt</i>	<i>nōlunt</i>

Over the course of early Latin, the differences between the two sets of variants broke down completely, and a distinction between emphatic and unemphatic variants was no longer made. Additionally, the SEMI-ANALYTIC forms with *ne* in the second person and third person singular were replaced across the board. These changes are further summarized in Table 3.29.⁴⁷

While the development of these forms is complex, there are two specific changes that

alternation would have had to been morphologized, making it more difficult to maintain *ne-* in only the second person and third singular. Instead, speakers could simply combine the regular negator with the verb. Other unverbated negated verbs do not show an intra-paradigmatic alternation; *ne* is found throughout. The presence of *ne-* would have been easier to maintain given that there was no alternation to uphold.

⁴⁶The exact nature of this availability is unclear. It is possible that while the forms with *nōn* were already the regular forms, the forms with *ne* were restricted to certain sociolinguistic contexts, such as register.

⁴⁷The additional change of *o > u* / ____ † is included.

Emphatic and Unemphatic

	SINGULAR	PLURAL
1ST	<i>nōlō</i>	<i>nōlumus</i>
2ND	<i>nōn vīs</i>	<i>nōn vultis</i>
3RD	<i>nōn vult</i>	<i>nōlunt</i>

Table 3.29: Final distribution of *nōlō* in Classical Latin

obscure the attested forms, i) the lack of contraction in the second person and third person singular in the unverbated paradigm (Stage 3), and ii) the replacement of *ne* with *nōn*, beginning in the pre-Plautine period. Otherwise, the proposed development follows directly from a combination of regular sound changes and uncontroversial analogical restorations (e.g. **no* with **ne*). The replacement of *ne* with *nōn* itself should also not be considered particularly contentious, and the distribution in Plautus follows from this assumption. Importantly, the replacement of the forms with *ne* with those with *nōn*, e.g. *nevīs* → *nōn vīs* is not dependent on any morphological or phonological rules. The old negative marker *ne* is simply replaced by the more productive *nōn*, which occurred in the vast majority of negative contexts.

The lack of contraction in the 3SG and 2PL, however, is more difficult to explain. Given the fact that the other forms contracted via regular phonological change, the non-contraction in the second plural **nouoltis* and third singular **nouolt* is remarkable and has so far been left unexplained by the broader literature on irregular Latin verbal morphology, from both diachronic and synchronic perspectives.

In my view, a proper explanation of the development of these forms needs to answer the following two questions:

1. What factors governed the choice of the ANALYTIC/SEMI-ANALYTIC forms in the 2PL and 3SG present indicative, while the other forms that could were contracted (the 1SG, 1PL, and 3PL forms)?⁴⁸

⁴⁸cf. Stage 3 as represented in (10)

2. How does the distribution of ANALYTIC and SEMI-ANALYTIC forms interact with the surprisingly similar distribution of athematicity in other irregular verbs?

Ultimately, there are two main approaches that can be taken with regard to determining why the 2PL and 3SG did not contract. The first is phonological; some factor prevented the contraction in these forms while allowing the contraction in the 1SG, 1PL, and 3PL, likely due to the difference in syllable structure. The second is morphological, in which the irregular distribution of other verbs like *ferō*, *sum*, etc. exerted analogical influence. The next few sections discuss the merits of these approaches.

3.3.2 PHONOLOGICAL APPROACH

A phonological approach to explaining the distribution should aim to unite the 2PL and 3SG via a unified phonological environment. Even though the verbal base / $\text{u}\text{o}\text{ł-}$ / was available for all person/number combinations aside from the 2SG (where we find / $\text{u}\text{i:}$ /), there is a phonological characteristic found only in the second plural and third singular, namely that the root syllable is a closed syllable (due to the athematicity of these forms). This contrasts with the other forms of the present indicative (apart from the second singular) that would have had open syllables, i.e. **nouolV*, **māuolV*. If contraction due to loss of -u- was blocked before a heavy syllable rather than an open syllable, then the alternation between SYNTHETIC and SEMI-ANALYTIC forms would simply be the phonologically expected outcome.

However, a specific restriction on contraction in heavy syllables is difficult to motivate based on the Latin evidence alone. Contraction explicitly due to deletion of -u- clearly preceded open syllables, e.g. *nōlō* < **novolō*, *nolumus* < **novolomos*, heavy syllables with one final consonant, e.g. **sei veis* ‘if you wish’ > *sīs*, and clusters that serve as the onset for the following syllable (i.e. *muta cum liquida*), e.g. *lavātrīna* ‘place for

bathing’ (Non. 212.7) > *lātrina*, *oblīvīscar* ‘I will forget’ > *oblīscar*. Syllables before other internal clusters also may contract. Consider the contraction that took place before the liquid *-r-* preceding a consonant: *altrōvorsum* ‘on the other side’ > *altrōrsum* (also *prōrsus* ‘straightforward’ < **prōvorsus*, *dextrōrsum* ‘towards the right side’ < *dextrōvorsum*).⁴⁹ There is little direct evidence that points towards a specific restriction before a *-lC-* cluster (besides *nōlō* and *mālō*) or even other clusters.

In addition to these contracted forms, there is a Plautine form that would support the phonological possibility that the 2PL should have contracted. The form *sultis* ‘if you (PL) want’, attested twelve times in Plautus, is seemingly contracted from *sī* ‘if’ + *voltis* ‘you (PL) want’. The form itself cannot be formally derived through a contraction, given that *-u-* would not have been lost between *ī* and *o*. Rather, as Sturtevant (1916) suggests, it may have been formed through analogy with *sīs* ‘if you (SG) want’ < **sei ueis*.⁵⁰ From a synchronic perspective, this form was clearly permissible, suggesting that the phonologically comparable form **nōltis* likewise should have been permitted, if not a hypothetical **noltis* with a short vowel.⁵¹

We may also consider the possibility that, for the 3SG form **nouolt* specifically, the monosyllabicity of its SYNTHETIC counterpart **nōlt* alongside its super heavy nature might have not been permitted by the phonological grammar of the time. Thus, if there was restriction that prevented the contraction from **novolt* to **nōlt* based on its syllabicity, we would likely expect the restored *nevolt* and eventually *nōn volt* as a repair.

⁴⁹Note that Plautus has uncontracted forms for *altrovorsum* (Cas. 555) and *dextrovorsum* (x3), but contracted *prōrsus* (x4) and *prōrsum* (x6). While it is mechanically possible that a potential **novoltis* would have remained uncontracted in Plautus, alongside the uncontracted forms of case in *altrovorsum* and *dextrovorsum*, it is likely that this would SEMI-ANALYTIC variant would have been leveled in favor of a SYNTHETIC version given the contracted forms elsewhere in the paradigm.

⁵⁰*vīs:sīs::voltis: X, X = sultis*.

⁵¹One may add to this the attestation of *nōltis* in Caecilius, discussed above in Section 3.2.2. Even though this form was discounted as a nonce formation, it could indicate that from a phonological perspective, it was permissible. I do not commit to any specific weight to this form in my analysis

Once again, however, the phonological evidence from Latin counters this hypothesis. As shown above, contraction caused by deletion clearly affected polysyllabic words, e.g. **novolō* > *nōlō* (three syllables contracting to two), **novolomus* > *nōlumus* (four syllables contracting to three). There is also at least one example of deletion of *-u-* creating a monosyllabic word after contraction: **sei veis* > *sīs* (Pl. *Aul.* 46, etc.). The form *Mārs* ‘Mars’ may be another example if contracted from the form *Māvors*.⁵²

Other contractions involving deletion of an intervocalic consonant were also able to produce monosyllabic words: for *-V_iV-* **trejēs* > *trēs*, for *-VhV-* *cohors* ‘farmyard, cohort’ > *cōrs*. This last example also indicates that contraction creating a monosyllable before a sequence of a liquid preceding a consonant was possible; contraction was permitted despite the fact that the syllable was super heavy. In all, monosyllables, super heavy or otherwise, created through contraction do not seem to have been restricted, making the lack of **nōlt* notable.

Beyond considering the potential impact of monosyllabicity on contraction, other phonotactic restrictions like prosodic minimality are not relevant here as monosyllabic words with heavy syllables are otherwise allowed (cf. *volt*). We can also exclude any influence from the morphological boundary, given the existence of the other forms in the paradigm that contract over the same morphological boundary (between **no-* and **vol-*) that is equally present in both the 3SG and 2PL.

In all, there do not seem to be any phonological restrictions on contraction that would have prevented the existence of the potential forms **nōlt* or **nōltis* that can be determined from our current understanding of the history of contraction in Latin. A phonological or phonetic restriction that specifically governed the sequence of *-olC-* in Latin would be *ad hoc*, given the lack of evidence, besides the absence of **nōlt*, **mālt*,

⁵²cf. Weiss (2020: 143).

etc. That said, such a restriction is not implausible *prima facie* and thus cannot be dismissed outright.

3.3.3 MORPHOLOGICAL (ANALOGICAL) APPROACH

Given that the outcome in the 2SG (*nevīs*) represents a phonologically regular outcome, an analogical treatment of the ANALYTIC/SEMI-ANALYTIC forms would first seek to explain the aberrant 3SG and 2PL forms in terms of a proportional analogy modeled on the 2SG. However, there is no obvious pathway for the introduction of *ne-* and *mā-* from the second singular to the other forms via a proportion. Apart from being athematic, the forms *volt* and *voltis* have little in common with *vīs*, making a direct comparison between the 2SG and 3SG/2PL difficult. A hypothetical proportion like *vīs* : *volt* :: *nevīs* : *X*, *X* = *nevolt* would be unmotivated, given that the sequence *volt* cannot be straightforwardly derived from *vīs*. The phonetic form of the 2SG simply does not predict the forms of the 3SG and 2PL.

Thus, a direct extension of *ne-/mā-* would require a significant amount of abstraction from the learner, which is not based on surface similarity. One possible factor to consider is the fact that the forms in the present tense of *nōlō* and *mālō* alternate in the same distribution as alternations in several other irregular verbs, as outlined in Section 3.2.4. Could these verbs have played a role in motivating an abstract analogy?

While it is difficult to devise a specific proportion, explaining the distribution via an abstract morphomic template is straightforward. The basic principle is schematized in what follows.

First, let us assume that there was no phonological restriction on contraction affecting **nouolt* and **nouoltis*, resulting in a 3SG **nōlt* and a 2PL **nōltis*. At this point, most irregular verbs would have had a templatic distribution as in (14-a), reflected in

terms of athematicity and suppletion. The ancestors of *nōlō* and *mālō* on the other hand would have alternated in terms of contraction in the distribution seen in (14-b), where only the 2SG is unexpected.

- (14) a. FER $\bar{O}$, SUM, ED $\bar{O}$, VOL $\bar{O}$: b. N $\bar{O}$ L $\bar{O}$ /M $\bar{A}$ L $\bar{O}$:

When the distributions of *nōlō* and *mālō* in (14-b) could no longer be synchronically derivable on phonological grounds, the distribution would pose a conundrum for learners. They would not only have to learn the irregular distribution in (14-a), but also the distribution in (14-b). On top of this, the forms in the 2SG, 2PL, and 3SG in *nōlō* and *mālō* would still have been athematic, further facilitating a connection between the irregular verbs and the derivatives of *volō*. In this sense, a learner would have been biased to extend the distribution from (14-a) to that in (14-b). This development is schematized below:

- (15) a. FER $\bar{O}$, SUM, ED $\bar{O}$, VOL $\bar{O}$: b. N $\bar{O}$ L $\bar{O}$ /M $\bar{A}$ L $\bar{O}$:

The prefixal element in *nōlō* and *mālō* was analogically restored on the basis of an attraction with the morphomic distribution affecting the other irregular verbs. The choice of which alternate (contracted versus uncontracted) to use in the 2PL and 3SG was determined on the basis of the second singular **no $\bar{\alpha}$ ueis*, which could not be contracted.

This proposal is not new; an abstract analogy of this type has already been suggested by Weiss (2020: 458), albeit without direct reference to a morphome: ‘The non-unverbated 2nd pl. may have been modeled on the 2nd sg. and the 3rd sg. was perhaps created on the pattern of other irregular verbs where the 2nd sg., 3rd sg., and the 2nd pl. cohere, e.g. *es, est, estis; fers, fert, fertis*.’ If we take into account that there is a set of verbs that engage in a generalizable pattern of athematicity in precisely this distribution, this analogical change does not seem implausible.

However, it should be noted that this account poses a problem for Maiden’s conception of a morphome. Recall that for the irregular verbs, athematicity and stem suppletion is not limited to the present indicative; it is also found in the present infinitive (e.g. *ferre, ēsse*) and the imperfect subjunctive (e.g. *ferrem, ēssem*). If this change was based on morphomic attraction in the sense of Maiden (2018), we would expect to find only uncontracted forms in the imperfect subjunctive and present infinitive. While we do find uncontracted forms in the imperfect subjunctive in Plautus (e.g. *Amph.* 512 *māvellem*), these are ultimately metrical variants, as I have argued for in Section 3.2.3. We find at least one contracted form in the imperfect subjunctive: *nōllem* (*Stich.* 513). In the Classical Period, these forms are always contracted. Additionally, we only find contracted infinitives for either these verbs (*nōlle, mālle*). In this sense, the attraction has only occurred in the present indicative, putting into question the coherence of the morphomic pattern.⁵³

The fact that only the present indicative of *nōlō* and *mālō* participate in the morphomic alternation puts into question whether the pattern in the irregular verbs should be classified as a morphome; I will return to this question in Section 3.4.1. However,

⁵³It should be noted that this is only strange from a morphomic perspective. A more traditional approach might not connect the imperfect subjunctive and present infinitive to the pattern in the present indicative.

for current purposes, this point is moot. Regardless of whether we call this change morphomic or not, it seems plausible that the irregular verbs might exert an analogical influence on *nōlō* and *mālō*, especially since *volō* also behaves in this manner.⁵⁴

3.3.4 WHAT NOW?

These two approaches leave admittedly leave us in a difficult position. While both present plausible pathways for the differentiation of SEMI-ANALYTIC and SYNTHETIC alternants in *nōlō* and *mālō*, there is little positive evidence for either approach.

For the phonological approach, a restriction on contraction into heavy syllables is phonologically plausible. However, the little data we have does not support such a restriction, given forms like *prōrsus* which involved contraction of *-ouo-* despite the following consonant cluster. On the other hand, we ultimately cannot dismiss the possibility that the scope of contraction was more limited before the attested period. The exact time frame of contraction in these sequences is notoriously difficult to determine (cf. Weiss 2020: 142-143). In this sense, the evidence from Old Latin and the Classical Period may not be fully informative with respect to a sound change with such a limited scope.

A morphological approach involving an abstract analogy between the irregular verbs and *nōlō*/*mālō* can be motivated mechanically, but is also not supported by any direct evidence. If, for example, the verb remained uncontracted in the present infinitive or imperfect subjunctive, it would support the notion of an abstract template affecting the morphomic distribution across the board.⁵⁵ Given the lack of evidence, it cannot be said for certain that the alternants were differentiated on the basis of an analogical

⁵⁴For a discussion of why other factors such as frequency and usage do not play into the distribution, see Appendix A.3.

⁵⁵However, one might also argue that this could support the notion of a phonological restriction on contraction, as these forms also contain heavy syllables.

change alone.

Thus, we can explain the differentiation in two ways:

Phonological Differentiation: The 2PL and 3SG did not undergo contraction due to a restriction based on the following heavy syllable, shown below:

(16) Phonological Differentiation

	SINGULAR	PLURAL		SINGULAR	PLURAL	
1ST	* <i>nouolō</i>	* <i>nouolomos</i>	⇒	1ST	* <i>nōlō</i>	* <i>nōlomos</i>
2ND	* <i>noueīs</i>	* <i>nouoltis</i>		2ND	* <i>noueīs</i>	* <i>nouoltis</i>
3RD	* <i>nouolt</i>	* <i>nouolont</i>		3RD	* <i>nouolt</i>	* <i>nōlont</i>

At this point, the restoration of **no-* with **ne-* targeted the 2SG, 2PL, and 3SG, allowing for the differentiation between the SYNTHETIC and SEMI-ANALYTIC forms.

It should be noted that after the alternation between the contracted and uncontracted forms became phonologically opaque, the alternation would have to be conditioned morphologically. As the distribution would be identical to that found in the present tense of the other irregular verbs, it is likely that the *nōlō* and *mālō* would be associated with the other irregular verbs from a synchronic perspective. In other words, the same morphological rules which governed the irregularity in *ferō*, *edō*, *etc.* would also be used to account for the SEMI-ANALYTIC forms in *nōlō* and *mālō*. In this sense, the phonological restriction would indirectly lead to a morphomic association, at least from a synchronic perspective.

Analogical Differentiation: Under this analysis, the 2PL and 3SG were subject to the contraction. However, they were analogically restored on the basis of the templatic pattern in the other irregular verbs:

(17) Analogical Differentiation

	SINGULAR	PLURAL		SINGULAR	PLURAL		SINGULAR	PLURAL
1ST	* <i>no<u>u</u>olō</i>	* <i>no<u>u</u>olomos</i>	⇒	1ST	* <i>nōlō</i>		1ST	* <i>nōlō</i>
					* <i>nōlomos</i>	⇒		* <i>nōlomos</i>
2ND	* <i>no<u>u</u>eis</i>	* <i>no<u>u</u>oltis</i>		2ND	* <i>no<u>u</u>eis</i>		2ND	* <i>no<u>u</u>eis</i>
					* <i>nōltis</i>			* <i>no<u>u</u>oltis</i>
3RD	* <i>no<u>u</u>olt</i>	* <i>no<u>u</u>olont</i>		3RD	* <i>nōlt</i>		3RD	* <i>no<u>u</u>olt</i>
					* <i>nōlont</i>			* <i>nōlont</i>

This analysis assumes that the uncontracted form in the 2SG served as the basis for the selection of the property which defines the morphomic distribution, namely the lack of contraction.

INTERIM CONCLUSION Above, I have just described two possible diachronic pathways for the development of *nōlō*. However, it should be noted that it is possible that both phonological and analogical pressures played a role in the development of the periphrastic alternation. For example, one might imagine a scenario in which there was a dispreference for contraction into a heavy syllable. However, this alone may not be enough to disallow the contraction. Given an alternative synchronic option, namely to extend the environment of the athematic distribution to the verb *nōlō*, learners may have resolved the problematic phonological sequence by associating the distribution in *NĀLĀ* with an already existing pattern. In this sense, both the phonological dispreference and morphological coherence conspired to produce the attested distribution.

Crucially, regardless of which approach one adopts, the distribution between SYNTHETIC and SEMI-ANALYTIC forms would have been associated with the distribution in the other irregulars. In what follows, we will discuss exactly what this relationship implies about the synchronic status of the morphomic distribution.

3.4 LATIN MORPHOME: SYNCHRONIC PERSPECTIVE

If analogy is part of the learning process, as assumed by many of the approaches discussed in Chapter 1, we would expect that a morphomic analogy as suggested in the previous sections was caused by the extension of a productive process (whether it be a template, pattern, or rule) to *nōlō*. In this sense, we should be able to describe the distribution in terms of a cohesive generalization that can be applied to new items. This section aims to investigate which frameworks best express the distribution and how this affects our understanding of morphological theory from a synchronic perspective.

There are two additional synchronic issues that this dissertation investigates with respect to non-root intra-paradigmatic alternations, i) the status of stems, and ii) the status of morphemes (cf. Section 2.5). The distribution in Latin involving the irregular forms outlined in Section 3.2.4 as well as the SEMI-ANALYTIC and ANALYTIC forms of *mālō* and *nōlō* broach both of these issues. At the root of this form is a pattern based on athematicity: an abstract morphological property which is reflected on the surface by the absence or presence of a theme vowel. As we will see, how we analyze athematicity can provide insights into both the status of stems and morphemes, as all of these aspects are intrinsic to the distribution in Latin.

This section expands on this idea by proposing a synchronic account that generates the distribution of the athematicity. Such an account is motivated by the coherence of idiosyncratic behavior across the set of a non-natural environments, i.e. the morphome.

I begin by justifying the morphomhood of the Latin pattern in Section 3.4.1. The bulk of the remaining section compares two theoretical frameworks, DISTRIBUTED MORPHOLOGY (Halle & Marantz, 1993) and PARADIGM FUNCTION MORPHOLOGY 2 (Stump, 2015), henceforth DM and PFM2 respectively, by analyzing the Latin pattern in each. I aim to demonstrate that PFM2 fails to provide a satisfactory explanation for the athe-

matic distribution; accordingly, I ultimately advocate for a piece-based framework like DM which is able to account for the athematic properties of these verbs.⁵⁶

3.4.1 AN ATHEMATIC MORPHOME

In this section, I argue that the distribution of athematic forms in the verbs *volō* ‘I want’, *edō* ‘I eat’, *sum* ‘I am’, and *ferō* ‘I bear’ constitute a morphome (Section 2.4). Under Maiden’s framing, morphomes can essentially be described as morphological templates which apply to a disjunct environment across a paradigm. Various morphophonological phenomena, e.g. suppletion, intra-paradigmatic alternations, etc. can be assigned to the cells that compose the template. This section discusses the pattern in Latin insofar as it relates to Maiden’s diagnostics for morphomhood.

Recall the two criteria for morphomhood presented in Section 2.4: a morphome is a distribution which i) cannot be explainable either phonologically or (morpho)-syntactically, and ii) shows coherence across several types of morphological expression (e.g. suppletion, affixation, periphrasis, apophony, etc.) In order to qualify for morphomic status, the following distribution in these Latin verbs needs to satisfy these criteria: the 2SG, 2PL, and 3SG of the PRESENT INDICATIVE, the PRESENT INFINITIVE, and the IMPERFECT SUBJUNCTIVE. I hope to demonstrate that these verbs meets the requirements in each of these environments.

The first criterion is satisfied in the Latin case. There is no syntactico-semantic feature or set of features that can account for the distribution that covers the second person (both singular and plural) and third person singular present indicative, imperfect subjunctive, and present infinitive. It may be possible to account for some parts of this distribution with syntactic features; however, the full distribution does not constitute

⁵⁶It may be possible to translate the proposed account to other piece-based frameworks.

a meaningful natural class. There are also no phonological factors that could unify the set of disjunct environments.⁵⁷ The verbs are not of similar phonological shape to one another, and some have direct minimal pairs that do not undergo the alternation, e.g. *ferre* versus *gerere*. In sum, this distribution perfectly aligns with the first criterion for morphomicity, given that the distribution must be defined in morphological terms.

The second criterion is more difficult to assess, but as we shall see, the pattern parallels other cases of morphemes, as multiple types of alternations can be shown to behave in the same manner. The full athematic behavior is summarized in Table 3.30:⁵⁸

⁵⁷At least for *ferō*, *edō*, *sum*, and *volō*. See Section 3.3.2 for a discussion of a potential phonological conditioning of *nōlō* and *mālō* in the PRES IND.

⁵⁸The table compares the six relevant verbs in terms of athematicity (ATH.), their athematic stems in the 2SG, 2PL, 3SG PRES IND, the PRES INF, and the IMPF, coherence (COH.), and finally their thematic stems in the 1SG, 1PL, 3PL PRES IND.

VERB	ATH.	2SG, 2PL, 3SG in PRES IND	PRES INF	IMPF SUBJ	COH.	NOTES:	1SG, 1PL, 3PL in PRES IND
SUM	✓	STEM= <i>es</i> <i>es, estis, est</i>	STEM= <i>es</i> <i>esse</i>	STEM= <i>es</i> <i>essem, etc...</i>	✓	Coherent Suppletion (<i>es-</i> vs <i>s-</i>)	STEM= <i>s(u)</i> <i>sum, sumus, sunt</i>
EDŌ	✓	STEM= <i>ēs</i> <i>ēs, ēstis, ēst</i>	STEM= <i>ēs</i> <i>ēsse</i>	STEM= <i>ēs</i> <i>ēssem, etc...</i>	✓	Coherent Suppletion (<i>ēs-</i> vs <i>ed-</i>)	STEM= <i>ed(i)</i> <i>edō, edimus, edunt</i>
FERŌ	✓	STEM= <i>fer</i> <i>fers, fertis, fert</i>	STEM= <i>fer</i> <i>ferre</i>	STEM= <i>fer</i> <i>ferrem, etc...</i>	✓	No Suppletion	STEM= <i>fer(i)</i> <i>ferō, ferimus, ferunt</i>
VOLŌ	✓	STEM= <i>vī/vol</i> <i>vīs, voltis, volt</i>	STEM= <i>vel</i> <i>velle</i>	STEM= <i>vel</i> <i>vellem, etc...</i>	✗	Non-coherent Suppletion (<i>vī-</i> vs <i>vol-</i> vs <i>vel-</i>)	STEM= <i>vol(u)</i> <i>volō, volumus, volunt</i>
NŌLŌ	✓	STEM= <i>nōn vī/nōn vol</i> <i>nōn vīs, nōn voltis, nōn volt</i>	STEM= <i>nōl</i> <i>nōlle</i>	STEM= <i>nōl</i> <i>nōllem, etc...</i>	✗	Periphrasis limited to PRES IND	STEM= <i>nōl(u)</i> <i>nōlō, nolumus, nōlunt</i>
MĀLŌ	✓	STEM= <i>māvi/māvol</i> <i>māvīs, māvoltis, māvolt*</i>	STEM= <i>māl</i> <i>mālle</i>	STEM= <i>māl</i> <i>māllem, etc...</i>	✗	Non-contraction limited to PRES IND	STEM= <i>māl(u)</i> <i>mālō, mālumus, mālunt</i>

Table 3.30: Summary of athematic behavior (Latin)

Each of these verbs has its own idiosyncratic behavior; however, I aim to show that athematicity is the property which truly unifies all of these verbs.

EDŌ ‘I EAT’: Beginning with the verb *edō* ‘I eat’, the morphomic coherence of athematicity is clear. The root allomorph *ēs* alternates with the allomorph *ed* throughout the non-perfective forms in the specific cells in question.⁵⁹ That is, the form *ēs* obligatorily co-occurs with athematicity.

SUM ‘I AM’: The verb *sum* mirrors the suppletive pattern of *edō*. The root allomorph *es* alternates with *su* in the present indicative and is present in the imperfect subjunctive and present infinitive. One difference between the distribution of suppletion of *sum* and *edō* is that the *es* allomorph is arguably present in the imperfect and future indicative, shown in Table 3.31, having undergone intervocalic rhotacism (*esV* > *erV*).

	1SG	2SG	3SG	1PL	2PL	3PL
IMPF IND	<i>eram</i>	<i>erās</i>	<i>erat</i>	<i>erāmus</i>	<i>erātis</i>	<i>erant</i>
FUT IND	<i>erō</i>	<i>eris</i>	<i>erit</i>	<i>erimus</i>	<i>eritis</i>	<i>erunt</i>

Table 3.31: IMPF and FUT IND of SUM in Latin

There are two possible analyses of the allomorph *er*. The first is to assume that, from a synchronic perspective, *er* is a morphological allomorph, distinct from an underlying /es-/. Alternatively, *er-* corresponds synchronically with an the underlying form *es-*, which means that the allomorph *es-* is not limited to the morphomic distribution. This may seem problematic; however, the forms in Table 3.31 are crucially not athematic. This means that we can state a generalization as ‘wherever *esse* is athematic, the exponent of the root is *es-*’. Thus, regardless of how we analyze *er*, the distribution of athematicity has a generalizable morphomic distribution.

⁵⁹The perfective forms have a third allomorph *ēd-*.

FERŌ ‘I BEAR’: The verb *ferō* replicates athematicity without making reference to suppletion.⁶⁰ Throughout the morphomic distribution, the root allomorph of *ferre* is *fer-*. However, as shown in Section 3.2.4.1, these very cells are obligatorily athematic, exemplified by the lack of a thematic vowel. As can be seen by the comparison with *gerō* in Table 3.20, Table 3.21, and Table 3.22, *ferō* lacks a theme vowel in the second person and third singular of the present indicative, the present infinitive, and the imperfect subjunctive. However, the rest of the paradigm parallels the thematic distribution shown by *gerō*. The verbs *ferō*, *edō*, and *sum* are thus similar in terms of how their morphological expression interacts with the specific cells in their respective paradigms. For *ferō*, the morphological expression is athematicity, while for *edō* and *sum*, athematicity co-occurs with suppletion. Crucially, even though the exact type of morphological expression differs the verbs, the distribution of athematicity is exactly the same.

VOĻŌ ‘I WANT’: The verb *volō* proves more difficult. Unlike the other verbs that undergo suppletion in the pattern (*edō*, *sum*), the root allomorphs of *volō* do not cohere across the different cells. In the present indicative, the allomorph *vol* is found in the third singular and second plural, while a highly irregular *vī* is found in the second singular. The allomorph *vel* is found in the imperfect subjunctive and present infinitive; however, it is also found in the present subjunctive (e.g. *velim*, etc). This means that *volō* does not have a suppletive alternation that follows the distribution of the morphomic pattern.

However, I contend there is still a meaningful pattern that can be deduced from the distribution. In the present indicative, even though the root allomorphy does not provide evidence of the aforementioned pattern, the second person and third person singular forms can be separated from the rest in terms of lack of a theme vowel, parallel

⁶⁰The verb *ferō* does have suppletion in its paradigm; however, these are not present in the *infectum*.

to the distribution of the theme vowel in *ferō*. In this sense, the special allomorphy of the second singular obscures the true abstract morphological alternation, thematicity versus athematicity. Likewise, even though the stem in the present infinitive and imperfect subjunctive is different from those in the present indicative, all cells are athematic. Regardless of the specific phonetic sequences that each allomorph exhibits, a morphological property, athematicity, accompanies all of the cells.

NŌLŌ ‘I NOT WANT’ AND MĀLŌ ‘I PREFER’: As discussed in Section 3.3.3, *nōlō* and *mālō* do not cohere in the present infinitive and imperfect subjunctive in terms of periphrasis and contraction respectively. This might suggest that they do not exemplify the same morphomic pattern as the other irregular verbs. However, it should be noted that, regardless of the SYNTHETIC-ANALYTIC distinction, the cells in the morphomic distribution are also athematic, a consequence of the fact that they are historically derived from *volō*. Thus, whatever governs athematicity in the other verbs from a synchronic perspective is likely to be responsible for athematicity in *nōlō* and *mālō*.

SUMMARY: In sum, each of the verbs demonstrates a slightly different pattern. Some of the verbs cohere completely in terms of their exponence; however, some verbs, like *volō*, show variation among the different cells. If we only take into account the stems, we might be inclined to say that these verbs are not morphomic. However, it is clear that there is a morphological property which unifies all of the verbs, namely athematicity. In this sense, the verbs are coherent across the morphomic distribution. I argue that this fulfills the second criteria for morphomehood.

3.4.1.1 THEMATICITY

Before contrasting how DM and PFM2 handle athematicity, there are a few points regarding thematicity that need to be discussed. Firstly, in traditional explanations, the alternation between *es* and *su* in the present indicative of *esse* cannot be expressed as an athematic alternation. However, I suggest it is worth considering whether the *u* in *su* has been reanalyzed as a theme vowel, at least synchronically. Certainly, in the plural forms, *sumus* and *sunt*, there is a striking parallel to the forms of *volō* for these cells, namely *volumus* and *volunt*, where the theme vowel synchronically can be described as a *-u-*. This allows for the present subjunctive and the present indicative of *esse* to be united under a singular root allomorph *s-*, with the marker *-i-* in the subjunctive to be related synchronically with the present subjunctive of *volō*. The assumption that the root allomorph *su* is indeed decomposable to *s-* and a theme vowel allows the alternation in the present indicative to further be abstracted as one that shows both athematicity and root suppletion, paralleling the other verbs that undergo an alternation in these cells.

A second issue pertains to whether the forms outside of the cells relevant to the morpheme are truly thematic. Let us consider, for example, *ferō* in more depth. This verb, along with other third conjugation verbs is characterized by a theme vowel *-e-*.⁶¹ However, unlike most first and second conjugation verbs, this theme vowel has a more limited distribution. Where we find first conjugation verbs with a future and imperfect indicative characterized by the *-ā-* theme vowel, e.g. *amābō* ‘I will love’, *amābam* ‘I was loving’, etc., the third conjugation has formations where the theme vowel is apparently missing, e.g. *geram* ‘I will bear’, *gerēbam* ‘I was bearing’. One might be tempted to

⁶¹One may describe the *-i-* that appears in the present indicative as an *-e-* that is synchronically raised to *-i-*.

say that these formations found in the third conjugation are athematic, therefore undermining the connection of *ferō* and *volō* with the other examples. However, I argue that this misses out on the crucial generalization. Comparing *ferō* with *gerō*, the forms for the future indicative, imperfect indicative, and present subjunctive are the same, the difference between the two verbs is characterized by a lack of *-e/i-* in the present indicative, present infinitive, and imperfect subjunctive. Whatever the reason there is no apparent theme vowel in the other forms, it is independent from the reason for the absence of the theme vowel in the morphomic pattern of *ferō*.⁶² The argument here boils down to the fact that *ferō* is athematic (as compared with other third conjugation forms) precisely in those cells where the theme vowel is normally present. Thus the athematicity that accompanies the morphomic distribution can be seen as a different ‘type’ of athematicity from that which is normal to the third conjugation.

A similar argument can be made to account for the present subjunctive of *sum*, *edō*, and *volō*, characterized by *-ī-*, originally from the old optative marker. The present subjunctive of the four morphomic verbs can be seen in Table 3.32 below. Arguably, none of the verbs have a theme vowel in these forms. However, as argued above, these forms are not athematic in the same way their imperfect subjunctive forms are, as other verbs in the third conjugation also lack theme vowels in these forms.

	FERŌ	VOLŌ	SUM	EDŌ
1SG	<i>feram</i>	<i>velim</i>	<i>sim</i>	<i>edim</i>
2SG	<i>ferās</i>	<i>velīs</i>	<i>sīs</i>	<i>edīs</i>
3SG	<i>ferat</i>	<i>velit</i>	<i>sit</i>	<i>edit</i>
1PL	<i>ferāmus</i>	<i>velīmus</i>	<i>sīmus</i>	<i>edīmus</i>
2PL	<i>ferātis</i>	<i>velītis</i>	<i>sītis</i>	<i>editis</i>
3PL	<i>ferant</i>	<i>velint</i>	<i>sint</i>	<i>edint</i>

Table 3.32: PRES SUBJ of FERŌ, VOLŌ, SUM, EDŌ in Latin

⁶²From a synchronic perspective, it is possible that the theme vowel is present in the cases where its absence is conspicuous. If we consider the fact that underlying short vowels are deleted before underlying long vowels, its absence is expected in these forms.

In this sense, the fact that there is no theme vowel in these forms is independent from the lack of theme vowel in the imperfect subjunctive.

In sum, I believe that the forms that do not occur in the scope of the morphomic pattern can all be shown to be analyzed as thematic, at least for the purposes of synchronic analysis.

3.4.2 DM VS PFM2: MOTIVATING A COMPARISON

The ultimate goal of modeling this distribution from a synchronic perspective is to formalize the morphomic pattern in a cohesive way. Doing so will motivate the extension of the pattern to novel items, as was the case for *nōlō*. If a framework cannot describe the generalization, there would be no reason for novel lexemes to change their morphological expression to cohere to the distribution. In what follows, I compare two frameworks, DISTRIBUTED MORPHOLOGY and PARADIGM FUNCTION MORPHOLOGY 2. As we will see, only DM can account for the generalization in a unified way.

This rest of this section outlines how the comparison between DM and PFM2 can further inform our understanding of both the status of stems and morphemes from a synchronic perspective.⁶³ As this section will illustrate, these frameworks differ in terms of how both of these concepts are handled. The subsequent analyses of athematicity in DM and PFM2 in Section 3.4.3 and Section 3.4.4 respectively demonstrates that a piece-based framework like DM, characterized by a rejection of both stems and explicit morphomic templates, is able to properly account for the generalizations in Latin. PFM2, however, is unable to capture the abstract property of athematicity, despite making explicit reference to both stems and morphomic properties.

At first glance, a DM-style account for the morphomic distribution faces an uphill

⁶³Related to this discussion is the comparison between DM and PFM2 in terms of theoretical restrictiveness in [Kramer \(2016: 109-116\)](#).

struggle. Recall that a morpheme is a morphological phenomenon whose distribution cannot be targeted by syntactic or phonological features alone.⁶⁴ The purported existence of a distinct morphological component directly counters the underlying assumption in DM that morphology as we observe it on the surface is actually ‘distributed’ across multiple grammatical components (Halle & Marantz, 1993), with features like *syntactic hierarchical structure all the way down* (Halle & Marantz, 1994) defining the basis of the framework. In this sense, a phenomenon, like a morpheme, which is dedicated explicitly to a morphological component of grammar explicitly challenges the underlying assumptions for DM.

The rejection of a separate morphological component in DM contrasts with the emphasis on morphology in PFM2. Indeed, in recent iterations of PARADIGM FUNCTION MORPHOLOGY, there has been emphasis on *morphomic properties*, which distinctly reference a morphological component (cf. Stump 2015, 2016). By comparing the two frameworks, I aim to show that a morphological component is not necessary to account for the highly abstract distribution in Latin.

With regard to the status of stems, one of the key differences between the two frameworks pertains to the existence of *morphemes* as distinct units.⁶⁵ In DM, words are built up through the combination of morphemes, the correspondence between terminal nodes in a syntactic hierarchy and the representations of their phonological exponence. PFM2, on the other hand, explicitly repudiates this notion and assumes a paradigmatic approach to explain the various morphological phenomena attested.⁶⁶ For our purposes, the important consequence of this assumption is that the building blocks of word formation are reduced to a choice of stem and various rules of exponence. These two building

⁶⁴Morphemes intrinsically are morphological in nature. Their existence is used as proof for an explicit morphological component of grammar (see Aronoff 1994).

⁶⁵Crucially, I refer to morphemes here, not *morphomes*.

⁶⁶This assumption is framed in the *irreducibility hypothesis* (Stump, 2015, 23).

blocks combine to form the individual cells in a paradigm. By comparing PFM2 with DM, we are essentially comparing *stem*-based and *morpheme*-based approaches.⁶⁷

As we will see, the pattern of athematicity in Latin uniquely touches on the comparison between these two approaches. Unlike other examples of morphemes, the Latin case exemplifies a distribution that contrasts an abstract distinction between athematicity and thematicity, which is not expressed through a one-to-one matching of stem type. As I aim to show, this *morphological* abstraction is difficult for PFM2 to handle, precisely because of its assumptions regarding the existence of stems. DM, on the other hand, is able to motivate the difference between athematic and thematic forms through the treatment of the thematic vowel as a morpheme. As we will see, the aggressive decomposition that a morpheme-based approach can take allows for the unification of the morphological phenomena seen in these irregular verbs.

To summarize, there are two ways in which a comparison between how the two frameworks account for the Latin case proves informative:

1. Do morphemes constitute evidence for the status of a distinct morphological component of grammar?
2. Is there a meaningful difference between how morpheme-based and stem-based approaches account for morphomic distributions?

Ultimately, I argue in favor of a DM-style approach, in the vein of recent work of Calabrese (2012, 2015b,a). This case study will show that the issues facing DM pertaining to morphomicity can be ultimately be overcome, as has been suggested by others (Embick, 2016, Trommer, 2016), providing support to the restrictive notion that attested morphological phenomena do not necessitate the existence of a distinct morphological component. Likewise, I aim to show that DM captures the generalizations of

⁶⁷See also the discussion in Section 2.5.1.

the phenomenon better than PFM2, specifically because a morpheme-based approach is used rather than a stem-based approach. Crucially, PFM2 fails to capture the abstract *morphological* property which unites the verbs in a morphomic pattern.

3.4.3 ATHEMATIC MORPHOME IN DM

Despite the common assumption that DM necessarily cannot account for morphomic phenomena, several attempts have been made to analyze morphomic distributions via impoverishment (Embick, 2016) or by feature insertion (Trommer, 2016). This section builds on the former strategy, in which the combination of impoverishment and vocabulary insertion produces the morphomic distribution.

3.4.3.1 MOTIVATING IMPOVERISHMENT

Frameworks that approach morphological phenomena in terms of an additive, piece-by-piece approach, e.g. DM (Halle & Marantz, 1993) or NANOSYNTAX (Starke, 2010), generally analyze theme vowels as a morpheme in and of themselves, although their exact morphosyntactic nature depends on the analysis. In the Latin data, all the forms involved in the distribution are unified by the fact that these forms lack a theme vowel, in opposition with the other forms of their respective paradigms. That is, all the verbs lack a theme vowel in the same disjunct set of environments. Despite the fact that suppletion seems to play a role in the alternation, I argue that it is the lack of a theme vowel that is at the heart of the alternation. Instead, the suppletion that patterns with athematicity is due to historical developments, and the fact that some of the forms explicitly do not have suppletion within the set of disjunct cells supports this notion.

In prior generative literature, thematic vowels have received much discussion (Oltra-Massuet, 2000, Oltra-Massuet & Arregi, 2005, Calabrese, 2015b, Fábregas, 2017, Pomino

& Remberger, 2022, *a.o.*), including Latin thematic vowels specifically (Embick, 2000, Embick & Halle, 2005). Analyses of theme vowels generally fall into two camps: i) one in which the thematic vowel itself does not have any specific syntactic content and is inserted *post-syntactically* (Oltra-Massuet, 2000, Oltra-Massuet & Arregi, 2005, Embick, 2010, *a.o.*), ii) the thematic vowel serves a specific syntactic role, such as a light verb (Fábregas, 2017) or as a verbalizer (Grestenberger, 2022). In terms of this analysis, the exact morphosyntactic implementation is not crucial and does not favor either one of these approaches. The intent of this section is to outline how the generalization is better accounted for using a morpheme-based analysis. Therefore, the framing of the proposal in this section should be seen as (relatively) theory-neutral and not as a specific commitment to the syntactic nature of the theme vowel slot/morpheme.

An analysis of the athematicity must account for the distribution, and I argue that the explicit nature of the theme vowel itself is almost irrelevant to its distribution. Rather, whatever governs its insertion (or lack thereof) is the important generalization that I aim to account for. I see two potential ways of expressing the notion of athematicity, i) either some ‘feature’ is deleted in the relevant environment, or ii) some ‘feature’ is inserted in the relevant environment. The first approach may be understood under a classical DM-style approach as *impoverishment* (or deletion) of the environment of the theme vowel. Essentially, whatever governs the insertion of the theme vowel is deleted in the athematic part of the paradigm. The second approach involves some type of feature insertion, rather than deletion. This could be arbitrary feature insertion similar to the analysis of morphemes by Trommer (2016) where redundancy rules allow for the insertion of an arbitrary feature, further feeding vocabulary insertion.⁶⁸

⁶⁸ Alternatively, one might posit some meaningful feature in the syntax that distinguishes some cells versus others; this would be akin to a Nanosyntactic approach. However, as we will see, a Nanosyntactic approach may have some of the same difficulties that a stem-based approach might have in terms of having homophonous stems express completely different structure/morphological information.

For the purposes of explaining the phenomenon in Latin, there is no reason to prefer feature deletion or feature insertion. That is, the specific Latin data do not strongly point to which type of analysis we should prefer. For theoretical purposes, I express the analysis in terms of feature deletion/impoverishment. The reason is two-fold. Firstly, the intuition behind the lack of a theme vowel suggests deletion is at play, rather than insertion. There is little reason, to my estimation, that a learner would posit feature insertion as an explanation for a lack of overt material. Secondly, allowing a morphosyntactic operation like redundancy rules potentially weakens the restrictiveness of the theory. In this case the problem is compounded by the fact that there is no clear morphosyntactic motivation for the insertion of a feature. That is, under an analysis in which a feature is inserted, it is likely that that feature would be arbitrary and otherwise unconnected with any morphosyntactic basis, as it is in the analysis of Romance in [Trommer \(2016\)](#). Admittedly, neither of these reasons is a knockdown for an analysis involving feature insertion; however, it is clear that if an analysis can handle the Latin case with feature deletion, it should be preferred over the alternative.

Moving forward to the proposal, I put forth the following intuition. There exist a certain number of features or conditions that cause the insertion of a theme vowel, $\{F_0 \dots F_n\}$. This depends on the framework, but I label this set of conditions with the variable X . In the cells in which a theme vowel is absent, X is deleted: $X \rightarrow \emptyset / __$ $\{2 \text{ PRES IND}\}, \{3\text{SG PRES IND}\}, \{\text{IMPF SUBJ}\}, \{\text{PRES INF}\}$. This impoverishment rule must be indexed for the set of verbs (*volō*, *ferō*, *sum*, *edō*) in which athematicity occurs. Crucially, the rule deletes whatever conditions the presence of a theme vowel and does not delete the theme vowel itself.

At this point, the nature of what X is can be fitted to suit various analyses that

allow impoverishment or otherwise feature deletion/restriction. In the next sections, I provide a plausible example for what the X could be by comparing this impoverishment rule to one found in a series of analyses for Italian athematic verbs in (Calabrese, 2012, 2015b,a).

3.4.3.2 ATHEMATIC VERBS IN ITALIAN

We may compare this impoverishment rule with a similar account for athematic verbs in Italian proposed in Calabrese (2012, 2015b,a). Here, Calabrese is dealing with cases in which regularly athematic verbs conjugate as if they are thematic in part of the paradigm. As an example, consider the forms of the perfect for some irregular Italian verbs, seen below in Table 3.33.⁶⁹ The highlighted cells show athematic forms, while the non-highlighted cells show regular thematic forms. This distribution is a remarkable parallel to the Latin irregular present indicative forms, although the two distributions are diachronically unrelated. Both distributions occur over a disjunct set of cells across a paradigm and involve an alternation of thematic versus athematic forms.

	1SG	2SG	3SG	1PL	2PL	3PL
VENIRE ‘come’	<i>vénni</i>	<i>venisti</i>	<i>véne</i>	<i>venimmo</i>	<i>veniste</i>	<i>vénnero</i>
METTERE ‘put’	<i>mìsi</i>	<i>mettèsti</i>	<i>mìse</i>	<i>mettémmo</i>	<i>mettèste</i>	<i>mìsero</i>
FÀRE ‘do’	<i>féci</i>	<i>facésti</i>	<i>féce</i>	<i>facémmo</i>	<i>facéste</i>	<i>fécero</i>

Table 3.33: Irregular PERF forms in Italian (adapted from Calabrese 2015b: 111)

To account for this distribution, Calabrese proposes that the roots that have a mix of thematic and athematic forms in the perfect are underlyingly marked with a [+TV-pruning] diacritic, where TV stands for theme vowel. These roots undergo impoverishment in the second singular and plural as well as the first plural. This proposal can be summarized in (18)-(19).

⁶⁹The forms in the table in Table 3.33 are retrieved from Calabrese (2015b: 111).

(18) $\text{Root}^{[+TV\text{-pruning}]} \rightarrow \text{impoverishment} \rightarrow \text{Root}$ (Calabrese 2015b: 112)

(19) Delete $[+TV\text{-pruning}]$ in the context $\{[+perf, +part, +pl]_{AGR}, [+perf, +part, -auth, -pl]_{AGR}\}$ (Calabrese 2015b: 112)

The basic idea behind this rule is that irregular athematic verbs are marked with a diacritic that marks the *pruning* (or deletion) of a theme vowel. Following the proposal in Calabrese (2012), impoverishment may target not only regular syntactic features but also diacritic features like a $[+TV\text{-pruning}]$. This impoverishment prevents the insertion of athematic forms, leading to the insertion of regular thematic forms, in accordance with a verb's lexical class. Calabrese exemplifies this by comparing the derivation of the perfect of the first singular and plural of the verb *scrivere* 'to write', namely *scrissi* 'I wrote' (athematic) and *scrivemmo* 'we wrote' (thematic). The two derivations can be seen in (20) and (21) for the first singular and first plural respectively.

(20) Derivation for *scrissi* 'I wrote' (Athematic) (Calabrese 2015b: 112)

- a. $[[[[\text{skriv}^{[+TV\text{-pruning}]}]_{\text{root}} \text{TV}] + \text{Perf TV}]_T + \text{part}, +\text{auth}, -\text{pl}]_{AGR}$
- b. $[[[[\text{skriv}^{[+TV\text{-pruning}]}]_{\text{root}}] + \text{Perf TV}]_T + \text{part}, +\text{auth}, -\text{pl}]_{AGR}$
- c. $[[[[\text{skris}]_{\text{root}}] \text{s}]_T \text{i}]_{AGR}$
- d. *scrissi* [skrissi]

(21) Derivation for *scrivemmo* 'we wrote' (Thematic) (Calabrese 2015b: 112)

- a. $[[[[\text{skriv}^{[+TV\text{-pruning}]}]_{\text{root}} \text{TV}] + \text{Perf TV}]_T + \text{part}, +\text{auth}, +\text{pl}]_{AGR}$
- b. $[[[[\text{skriv}]_{\text{root}} \text{TV}] + \text{Perf TV}]_T + \text{part}, +\text{auth}, +\text{pl}]_{AGR}$
- c. $[[[[\text{skriv}]_{\text{root}} \text{e}] \emptyset]_T \text{mmo}]_{AGR}$
- d. *scrivemmo* [skrivemmo]

The differentiation between the two derivations takes place in the second step. In (20-b), the [+TV-pruning] diacritic is not impoverished, as being a first singular form does not trigger the impoverishment. Therefore, the theme vowel *is* pruned. In (21-b), however, the fact that the form is a first plural (+part, +auth, +pl) triggers the impoverishment rule. Thus, the diacritic conditioning the loss of a theme vowel is deleted, causing the insertion of a theme vowel. The alternation between thematic and athematic forms is therefore created through the application or non-application of an impoverishment rule.

3.4.3.3 APPLICATION TO LATIN ATHEMATICITY

The same logic can be applied to the Latin case, albeit with some minor differences. In order to encode the general intuition that the theme vowel is deleted, we can restate (18) and (19) as below in (22) and (23) respectively.

$$(22) \quad \text{Root}^{[+TV]} \rightarrow_{\text{impov.}} \text{Root}$$

$$(23) \quad \text{Delete } [+TV] \text{ in the context } \{ [+pres, +ind, -part, -pl]_{\text{AGR}}, [+pres, +ind, +part, -auth]_{\text{AGR}}, [+impf, +subj]_{\text{T}}, [+pres, +inf]_{\text{T}} \}$$

The differences between this analysis and that of Calabrese (2012, 2015b,a) are minimal. For example, the target of the impoverishment rule is a diacritic that describes the theme vowel insertion rather than one that describes the pruning of a theme vowel. However, this does not present a big change, as the two diacritics can be seen as theoretical equivalents.⁷⁰

⁷⁰The diacritics in (22) and (23) could be written as [+TV-pruning]. The analysis would change slightly as the emphasis would be on the thematic cells in the paradigm rather than the athematic cells. In order to compensate, the contexts in (23) would need to change to those cells where the theme vowel was present, as those would be the contexts in which the diacritic was impoverished. This can be illustrated by simply changing the impoverishment rule as in (i).

$$(i) \quad \text{Root}^{[+TV\text{-pruning}]} \rightarrow_{\text{impov.}} \text{Root}$$

The exact conditions of the environment for an edited version of (23) depends on one's assumptions regarding the presence of the theme vowel in the other tense/mood combinations. For example, the imperfect indicative, future indicative, and present subjunctive in the third conjugation do not overtly show a theme vowel, although it may be present underlyingly; an underlying short vowel is deleted before a long vowel in Latin. The potential for deletion makes these tense/mood combinations ambiguous with regards to whether a thematic vowel is present underlyingly.

One could posit the following impoverishment rule that deletes a [+TV-pruning] diacritic in the present indicative, leading to the insertion of a thematic vowel in the 1SG, 1PL, 3PL.

- (24) Delete [+TV-pruning] in the context $\{[+pres, +ind, +part, +auth]_{AGR}, [+pres, +ind, -part, +pl]_{AGR}, \text{etc.}\}$

By using an analysis in the vein of the work of Calabrese (2012, 2015b,a), we have accounted for deletion of a theme vowel in the relevant forms of the irregular verbs in Latin as discussed previously. The deletion is achieved by means of an impoverishment rule, which can be written in various ways, depending on which theoretical assumptions one wants to make. The impoverishment rule itself targets the deletion of whatever inserts the theme vowel across the disjunct set of environments that show athematicity: $\{3SG\ PRES\ IND, 2SG\ PRES\ IND, 2PL\ PRES\ IND, IMPF\ SUBJ, PRES\ INF\}$.

There are several benefits to listing the morphomic distribution in this way. Most importantly, as the distribution is listed out as a rule, it can easily be extended to new verbs; we have to merely extend the set of verbs to which the rule applies to include the new lexical items. The second benefit is that the rule is listed in several disjunct

environments. Recall that the periphrastic alternation affecting *nōlō* is only found in the present indicative, not the present infinitive or imperfect subjunctive. This suggests that whatever rule or pattern was extended to *nōlō* must have been able to reference only part of the athematic distribution. While the presence of disjunct rules weakens our theory, it is unclear to me how to account for the Latin data otherwise.

3.4.3.4 VOCABULARY INSERTION

The second part of this analysis focuses on how the distribution of different root allomorphs within the irregular verbs follows directly from the Vocabulary items relevant to the specific verbs. This section assumes the general mechanics involved with Vocabulary insertion as laid out in [Halle & Marantz \(1993\)](#). I do, however, assume that the vocabulary items responsible for the different exponents of the roots participate in competition with one another (contra [Marantz 1997](#), [Harley & Noyer 1999](#), but with [Bobaljik 2012](#), [Harley 2014](#)). Beyond these assumptions regarding the process of Vocabulary Insertion, for purposes of clarity, I assume that Latin verbs have a general underlying structure; thus, I posit a simple set of features responsible for the exponence of the different tense and mood combinations for the spell-out of the individual Vocabulary items; however, these are framed generally and may be replaced with more specific claims in conjunction with other analyses (e.g. [Embick 2010, 2016](#)).

This section will outline the vocabulary items for *ferō* and *volō*. As we will see, these two verbs differ wildly in the expression of the allomorph of the root. Accordingly, there is little in common with the distribution of the respective exponents. However, this is unproblematic for a DM approach, as it is the impoverishment rule which accounts for the generalization that these verbs all show athematicity in the same set of forms. In this sense, there is no need for the individual vocabulary items to express the distribution.

The vocabulary insertion rules for the two verbs described in this section, *ferō* and *volō*, demonstrate how the the exponence of the root (or stem) is distinct from the athematicity. The vocabulary items for the other verbs can be found in Appendix A.4.

VI RULES FOR *FERŌ*: We may begin with the relevant spell-out rules for *ferre* ‘to bear’, as it is the most simple. As described in Section 3.2.4, the verb *ferre* does not exhibit root allomorphy within the context of the patterned athematicity. Instead, we find the sequence *fer-* throughout the *inflectum*. Thus, we can simply specify the allomorph *tul* for the perfective tenses and an elsewhere case expressed by *fer*, as presented in (25).⁷¹

(25) VOCABULARY ITEMS for *ferō* ‘I bear’

- a. $v_{\text{bear}} \leftrightarrow \text{tul} / \text{ ___ } [+perf]$
- b. $v_{\text{bear}} \leftrightarrow \text{fer}$

VI RULES FOR *VOLŌ*: The next set of Vocabulary items to examine are those of *velle* ‘to want’, represented in (26).

(26) VOCABULARY ITEMS for *volō* ‘I want’

- a. $v_{\text{want}} \leftrightarrow \text{vel} / \text{ ___ } [-perf, +subj], [+pres, +infinite]$
- b. $v_{\text{want}} \leftrightarrow \text{vī} / \text{ ___ } [+2, +sg, +pres, +ind]$
- c. $v_{\text{want}} \leftrightarrow \text{volu} / \text{ ___ } [+perf]$
- d. $v_{\text{want}} \leftrightarrow \text{vol}$

⁷¹I ignore here and in the following lexical entries the notion of a *third stem*, as introduced by Aronoff (1994). The *third stem* refers to the stem adhered to by the perfect and future participles and has been the subject of much debate in the theoretical literature Aronoff (1994), Embick & Halle (2005), Steriade (2016), Kodner (2022), *a.o.* This stem affects our data in the sense that some of the lexical entries actually require more entries, in order to account for the third stem (e.g. the participial stem *lāt-* for *ferō*). However, these rules do not otherwise affect the pattern of athematicity.

The root of this verb shows several allomorphs, none of which are correlated with the patterned athematicity. The allomorph *vel*, the most faithful to its prehistorical source **uelh₁-*, is seen in the present subjunctive, imperfect subjunctive, and present indicative. As the present infinitive does not share a feature set as the subjunctive forms, it is necessary to posit a disjunctive Vocabulary Item. Next, the 2SG PRES IND shows the allomorph *vī*, present only in this cell of the paradigm. I also include an allomorph for the perfective forms, the exponent of which is *volu*; this form generally parallels the general distinction between the perfective forms and the imperfective forms expressed in many verbs. However, it should be noted that it is also possible to segment *volu* as *vol-u-*, where the *-u-* represents a suffix *-u-* (orthographic <v>), which is the perfective marker seen in most verbs, e.g. *am-ā-v-ī* ‘I loved’, and is thus distinct from the thematic vowel characteristic of the third conjugation. Finally, the form *vol* that is found throughout the rest of the paradigm is the elsewhere case, including all forms of the present indicative besides the 2SG.⁷²

SUMMARY OF VI RULES: The vocabulary items for these two verbs show distinct environments. It is clear that the Vocabulary Insertion rules do not correlate with the athematic distribution.⁷³ Instead, Section 3.4.3.3 has shown that the irregular behavior of these verbs combined with the pattern of athematicity can be accounted for by an impoverishment rule that combines with a set of individual spell-out rules for the different verbs in question. While the spell-out conditions for the different allomorphs

⁷²There is a question of whether the distribution for the allomorphs *vel* and *vol* for *velle* may be handled by readjustment rules rather than by a contextual allomorphic account. The practical consequence of proposing a readjustment rule instead of a competitive vocabulary item is that it may eliminate one of the vocabulary items; however, instead, one must posit a readjustment rule with the same environment as the vocabulary item. As far as I can see, there is no immediate reason to assume either analysis over the other.

⁷³See Appendix A.4 for the VIs of *esse*. The exponent *ēs* does correspond to the athematic pattern. Regardless, it is clear that the verbs as a group do not exhibit cohesion in terms of their root exponence.

of the verbs are not unified, this follows from the fact that we are discussing irregular verbs. These verbs are specifically characterized by their idiosyncrasy and much of their exponence must simply be memorized. The effect of the singular impoverishment rule is not to deny the status of these verbs as irregular; rather, the rule is the result of a generalization over a set of irregular data, demonstrating that learners still try to find regularity among irregular words.

3.4.3.5 DISCUSSION OF DM APPROACH

The DM approach discussed above has a clear pathway for the generalization to be extended to novel items. The athematic environment is formalized cohesively as an impoverishment rule, while the exponence of the individual irregular lexical items is listed in a separate component. The environment can be extended to other verbs by simply increasing the set of verbs the rule applies to.

The benefit of using a morpheme-based (or piece-based) analysis is two-fold. The first benefit is that it captures the generalization that the presence of the theme vowel alternates with its absence in the present indicative, imperfect subjunctive, and the present infinitive. Particularly illuminating is the present indicative of *ferō*, where the same ‘stem’ *fer-* appears in both the thematic and athematic forms. The only difference in the exponence between these forms is whether the thematic vowel is present or not, indicating that the theme vowel is segmentable in its own right; it is the theme vowel itself which participates in the alternation, not the exponent of the root.

This analysis of the theme vowel contradicts the segmentation of Latin verbs in most stem-based analyses (cf. [Lieber 1980](#), [Aronoff 1994](#), [Stump 2015](#)), which are unable to account for the difference between the ‘stems’ with *fer-* and *fere-* convincingly. As we will see in the analysis in PFM2, stem-based analyses are ultimately committed to positing

two different stems for *ferō* that happen to look identical apart from the presence or absence of a final vowel.⁷⁴

The second benefit of a morpheme-based analysis is that it allows for the separation between what governs root allomorphy and what governs athematicity. This benefit is exemplified specifically by the distribution of athematic forms of *volō*, which do not show the same root exponent in every context. For the PRES IND 2SG the form is *vī*, while the rest of the present indicative, including both thematic and athematic forms has *vol*. The imperfect subjunctive and present infinitive exhibit yet another exponent, *vel*. In order to account for all these irregular forms together, we must look at the athematicity, not the exponents of the root. The forms of *volō* in the 2SG, 3SG, and 2PL PRES IND, the IMPF SUBJ, and the PRES INF can only be unified if the alternation is stated in terms of the presence of absence of the theme vowel.

If the stem is the comparandum, rather than athematicity, we cannot describe the alternation in *volō* as a unified phenomenon. This in turn means that the fact that *volō* has a pattern of athematicity in exactly the same environments as where *edō* is athematic is simply a coincidence. In other words, relying on the stem to describe the alternation restricts the scope of the generalization as fewer verbs participate.

To summarize, a morpheme-based account can describe the distribution of thematic and athematic forms in terms of the presence and absence of the theme vowel. I have argued that this type of approach is able to account for the generalizations seen in the Latin data. As we will see, PFM2 has difficulty describing the same generalizations in a united way. Accordingly, the framework cannot motivate the extension of the rule or the environment to other lexical items.

⁷⁴In PFM2, one might imagine that a *stem-formation* rule might be used to derive the thematic stem from the athematic stem. This would work mechanically, but in the context of the fact that suppletion often correlates with the athematicity, an account in which *fere-* is formed by a stem-formation rule, but *ede-* is not is unconvincing.

3.4.4 ATHEMATIC MORPHOME IN PFM

3.4.4.1 BRIEF BACKGROUND ON PFM2

In this section, I provide an analysis for the Latin data in PFM2 using the mechanics illustrated in [Stump \(2015\)](#).⁷⁵ For our purposes, we can limit the analysis to the use of four concepts. The first is the **Stem** function, which assigns the base stem to the different cells in a paradigm. As we will see, the bulk of the exponence of the different verbs is handled directly by this function. The second relevant mechanic is the **pm** (property mapping) function. In PFM2, different cells are associated with different property sets. The **pm** function allows the assignment of non-syntactic features like declension class or morphomic properties to the property set of a specific cell. The **Corr** function applies these two functions to the paradigm function for the cell to convert the content cell to the form cell. Finally, *rules of exponence* play a minor role in the discussion of theme vowels. Rules of exponence are rules which govern the addition of affixes to a stem and are assigned in different *blocks* whose order and content is determined on a language-specific basis.

To illustrate how these mechanics are utilized, I outline the derivation of the Latin 2SG PRES IND *parās* and the 2SG IMPF IND *parābās* of the first conjugation verb *parō* ‘I prepare’ in Table 3.34 as described in [Stump \(2015\)](#).

The starting point for the derivation is the paradigm functions for the specific cells which denote the content of the cell. The functions for the two cells are below:⁷⁶

- (27) a. PF(⟨PARĀRE, {2sg pres ipfv act ind}⟩)
b. PF(⟨PARĀRE, {2sg pret ipfv act ind}⟩)

⁷⁵For the purposes of brevity, I assume that the reader has familiarity with the framework. For more in-depth discussion, see [Stump \(2015\)](#), [Bonami & Stump \(2016\)](#).

⁷⁶*ipfv* stands for imperfective.

		IMPERFECTIVE	PERFECTIVE
1SG	PRES	<i>parō</i>	<i>parāvī</i>
2SG		<i>parās</i>	<i>parāvistī</i>
3SG		<i>parat</i>	<i>parāvit</i>
1SG	FUT	<i>parābō</i>	<i>parāverō</i>
2SG		<i>parābis</i>	<i>parāverīs</i>
3SG		<i>parābit</i>	<i>parāverit</i>
1SG	PRET	<i>parābam</i>	<i>parāveram</i>
2SG		<i>parābās</i>	<i>parāverās</i>
3SG		<i>parābat</i>	<i>parāverat</i>

Table 3.34: *parō* ‘I prepare’ in Latin (Stump 2015: 77)

These two functions represent the content cells, where the lexeme is associated with the specific syntactic properties of the cell. The next step is to convert the content cell to a form cell by means of the **Corr** function, where **Corr**($\langle L, \sigma \rangle$) is the *form correspondent* (Stump, 2015) of the content cell. The values of this function are in turn determined by the **Stem** and **pm** functions. Thus, **Corr**($\langle L, \sigma \rangle$) = $\langle \mathbf{Stem}(\langle L, \sigma \rangle), \mathbf{pm}(\sigma) \rangle$.⁷⁷

For the two cells, there are several stems to choose from as per Stump (2015: 118): i) the imperfective stem *parā*, ii) the perfective stem *parāv*, and iii) the *b* stem *parāb*. For the two cells, the imperfective stem is used for the 2SG PRES IND ACT and the *b*-stem for the 2SG IMPF IND ACT. The **Stem** function assigns the correct stems to the form cell by means of the **Corr** function.

As for the **pm** function, it assigns non-syntactic features to the property set. In this case, we only need to assign the conjugation class. For PARĀRE, the **pm** function assigns the *conj1* feature, as it is a first conjugation verb.

The application of both of these functions through use of the **Corr** function produces the following output:

⁷⁷The **Stem** function may also reference the **pm** function.

- (28) a. $\text{PF}(\mathbf{Corr}(\langle \text{PAR}\bar{\text{A}}\text{RE}, \{2\text{sg pres ipfv act ind}\} \rangle)) = \text{PF}(\langle \text{par}\bar{\text{a}}, \{2\text{sg pres ipfv act ind conj1}\} \rangle)$
 b. $\text{PF}(\mathbf{Corr}(\langle \text{PAR}\bar{\text{A}}\text{RE}, \{2\text{sg pret ipfv act ind}\} \rangle)) = \text{PF}(\langle \text{par}\bar{\text{a}}\text{b}, \{2\text{sg pret ipfv act ind conj1}\} \rangle)$

Next, the form cell is converted to its corresponding realized cell. At this point, rules of exponence apply in blocks, leading to the addition of inflectional affixes. This occurs in different blocks which are specified for certain rules of exponence. To illustrate this, the addition of the person/number marking for the person and number markers *-s* and *ās* is shown in (29-a)-(29-c).

- (29) a. **Block 1:**
 $\text{X}, \sigma : \{2\text{sg}\} \rightarrow \text{Xs}$
 $\text{X}, \sigma : \{2\text{sg pret ipfv}\} \rightarrow \text{X}\bar{\text{a}}\text{s}$
 b. $\text{PF}(\langle \text{par}\bar{\text{a}}, \{2\text{sg pres ipfv act ind conj1}\} \rangle) = [\text{I} : \langle \text{par}\bar{\text{a}}, \{2\text{sg pres ipfv act ind conj1}\} \rangle] = \text{par}\bar{\text{a}}\text{s}$
 c. $\text{PF}(\langle \text{par}\bar{\text{a}}\text{b}, \{2\text{sg pret ipfv act ind conj1}\} \rangle) = [\text{I} : \langle \text{par}\bar{\text{a}}\text{b}, \{2\text{sg pret ipfv act ind conj1}\} \rangle] = \text{par}\bar{\text{a}}\text{b}\bar{\text{a}}\text{s}$

(29-a) refers to the block that contains the rules of exponence that governs the addition of the person/number suffixes. These rules, once applied to the form cells, produce the final realization *parās* and *parābās*, as seen in (29-b) and (29-c) respectively.⁷⁸

To summarize, the cells that are realized as *parās* and *parābās* are the result of several functions that apply to the content paradigm and form paradigm. The **Stem** function determines the main choice of stem, while the **pm** function assigns morpho-

⁷⁸Conceivably, one might analyze the vowel in the ending *-ās* as another suffix that should have its own rule block.

logical properties to the cells, including conjugation class and morphomic properties. These two functions apply to the **Corr** function to produce the form cell. Finally, the rule blocks determine the exponence of affixes.

FURTHER ASSUMPTIONS For the purposes of clarity, I simplify the transition from the form cell to the realization by putting aside how the verbal suffixes that mark tense, mood, and person/number are exponed, as these suffixes can be handled via their respective blocks. Ultimately morphomicity and athematicity are unconnected to the spell-out of these suffixes. There is, however, one exception, as the theme vowel is inherently connected to the question of athematicity. Crucially, in this analysis I do not decompose the theme vowel into its own morphological block, following the outline of stem types for Latin as set out in [Stump \(2015\)](#).⁷⁹ This segmentation is also in agreement with the STEM MAXIMIZATION PRINCIPLE [Taylor \(2007\)](#).⁸⁰

Encoding the theme vowel directly into the stem has an important consequence for this analysis: there is no meaningful structural difference between thematic and athematic stems. Thematic stems simply have one of several vowels appended to the end of their stems, while athematic forms simply lack any vowel at the end of their stem. A stem-based analysis therefore differs from a morpheme-based analysis in that thematicity (or lack thereof) is not directly encoded at any point in the derivation. To prefigure the analysis of *ferō*, the stems *feri* (1PL *ferimus*) and *feru* (3PL *ferunt*) under a PFM analysis are not thematic in opposition with an athematic stem *fer*. Rather, *feri* and *feru* are simply two of the many stems listed in the **Stem** function of *ferre*.⁸¹

The lack of a distinction between athematic and thematic stems also means that

⁷⁹I consider the alternative in which the theme vowel is decomposed below.

⁸⁰This principle determines that ‘if a (partial) word form is used for more than one cell in a lexeme’s paradigm, treat it as a stem’ ([Taylor 2007](#): 231)

⁸¹In other words, *feru* is just as unrelated to *feri* as either is to *feri* or *tul*.

stems that lack a theme vowel are not united as an athematic class, as we will see shortly with the distribution of stem alternates of the verb *volō*.⁸² Recall that this verb has three root allomorphs that show athematic outcomes, *vī*, *vol*, and *vel*. Accounting for these forms together is not possible using one singular **Stem** function. For this reason, it is also not possible to encode a specific conjugation class just for the stems that show athematicity into the **pm** function.

Otherwise, the general choice of stem formation and property mapping in this analysis is modeled after the stems proposed in the analysis of the Latin verbal system in [Stump \(2015\)](#).

3.4.4.2 STEM FUNCTIONS

This section will examine the stem functions for *ferō* and *volō*, the same verbs discussed in the DM analysis in Section 3.4.3.4. The stem functions for the other verbs can be found in Appendix A.5. This section demonstrates that the stem functions for the different verbs do not cohesively refer to the athematic distribution.⁸³ Thus, the stem functions do not adequately provide a unified explanation for the athematic generalization, the main goal of a synchronic explanation of the distribution.

⁸²This is not to say that a general distinction between different types of conjugation class, including thematic and athematic types is impossible in PFM2.

Let us say there were two types of verbs, i) athematic verbs and ii) thematic verbs with a theme vowel that occurs throughout the paradigm. In such a case, we could simply apply a conjugation feature via a property mapping which would distinguish between the two types. This is essentially the situation of athematic and thematic stems in Ancient Greek. Here, athematic and thematic types are differentiated on the basis of i) the presence or absence of a theme vowel, and ii) the inflectional ending in the 1SG PRES IND (*-ō* in thematic stems, *-mi* in athematic stems). In PFM2, we could apply a property *athematic* to verbs which not only lack a theme vowel but also take *-mi* in the first person. Similar also is the analysis of Sanskrit verbs throughout [Stump \(2015\)](#) (cf. the property **root** of root aorists).

In Latin, athematicity plays a slightly different role. The general distinction between thematic and athematic verbs has largely eroded. Instead, athematicity is preserved in certain irregular verbs in the morphomic distribution (2SG, 2PL, 3SG PRES IND, IMPF SUBJ, PRES INF). Here, athematicity is a property which does not

⁸³Apart from the verb *edō*. See Appendix A.5.1.

STEM FUNCTIONS FOR THE VERB FER $\bar{O}$: The different stem functions associated with the final verb FERRE ‘to bear’ are given below in (30).

(30) Clauses in the **Stem** function in Latin relevant to FER $\bar{O}$ ‘to bear’

- a. **Stem**($\langle \text{FERRE}, \sigma : \{\text{perfective}\} \rangle$) = *tul*
- b. **Stem**($\langle \text{FERRE}, \sigma : \{\text{impf ind}\} \rangle$) = *ferēb*
- c. **Stem**($\langle \text{FERRE}, \sigma : \{\text{pres ind 1 pl}\} \rangle$) = *feri*
- d. **Stem**($\langle \text{FERRE}, \sigma : \{\text{pres ind 3 pl}\} \rangle$) = *feru*
- e. **Stem**($\langle \text{FERRE}, \sigma : \{\} \rangle$) = *fer*

There is a clear distinction between the perfective stem and the other forms. The forms of the imperfective, however, have to be broken down further into several stems. The *b*-stem corresponds with other imperfect stems (see below for VOL $\bar{O}$). The present indicative itself must be broken down into three different stems: *feri* which corresponds to the 1PL *feri-mus*, *feru* which corresponds to the 3PL *feru-nt*, and *fer* which corresponds to the rest of the present indicative, including the athematic forms.⁸⁴

Crucially, none of the environments of the stem functions refer to the athematic distribution. Rather, the athematicity is expressed by the stem *fer* which happens to arbitrarily lack a stem-final vowel.

STEM FUNCTIONS FOR THE VERB VOL $\bar{O}$: The stems, as seen in (31), pertaining to the verb *volō* are shown below:

(31) Clauses in the **Stem** function in Latin relevant to VOL $\bar{O}$ ‘to want’

- a. **Stem**($\langle \text{VELLE}, \sigma : [\{\text{pres ind}\} \wedge [\{1 \text{ pl}\} \vee \{3 \text{ pl}\}]] \rangle$) = *volu*

⁸⁴The imperfect subjunctive stem could be analyzed as *ferrē*, cf. fn. 9 in Appendix A.5.1.

- b. **Stem**($\langle \text{VELLE}, \sigma : \{\text{impf ind}\} \rangle$) = *volēb*
- c. **Stem**($\langle \text{VELLE}, \sigma : \{\text{pres ind 2 sg}\} \rangle$) = *vī*
- d. **Stem**($\langle \text{VELLE}, \sigma : [\{\text{imperfective subj}\} \vee \{\text{pres inf}\}] \rangle$) = *vel*
- e. **Stem**($\langle \text{VELLE}, \sigma : \{\text{perfective}\} \rangle$) = *volui*
- f. **Stem**($\langle \text{VELLE}, \sigma : \{\} \rangle$) = *vol*

In of themselves, the stems are unproblematic. The imperfective subjunctives (including both the PRES SUBJ and IMPF SUBJ) along with the present infinitive correspond to the realization *vel*.⁸⁵ The perfective stem is realized uncontroversially as *volui*, while the imperfect stem *volēb* corresponds to the *b*-stem that we have already seen for *ĒSSE* (i.e. *edēb*).

With regards to the present indicative, PRES IND 2SG corresponds to the form *vī*, while the 1PL *volumus* and 3PL *volunt* imply the stem *volu-*. Assuming this, the other forms of the present indicative (1SG *volō*, 3SG *volt*, and 2PL *voltis*) correspond to the elsewhere case which is realized as *vol*.

Once again, the distributions account for the forms, but cannot express the abstract morphological notion of athematicity. What is crucial for the generalization is the relationship of the stems *vī-*, *vol*, and *vel*. Under this analysis, these three stems are unrelated, making it impossible to account for the athematicity using just stems. Accordingly, there is also no unified reference to the irregular pattern described in Table 3.30.

SUMMARY Neither of the stem functions for the two verbs discussed above reference the athematic distribution. Instead, the stem functions have idiosyncratic distributions, similar to the Vocabulary Items described in Section 3.4.3.4.

⁸⁵One could also analyze the present subjunctive with a special stem *velī*. Additionally, the imperfect subjunctive could be analyzed with a stem *vellē*.

3.4.5 THEME VOWELS

The stem functions for the aforementioned verbs can all mechanically describe their respective realized forms. However, the verbs cannot be unified in terms of either athematicity or a morphomic distribution. One may wonder whether decomposing the stems would allow for a more unified analysis. This section considers that possibility. Specifically, I consider an approach in which the generalization is expressed in terms of an inflectional class which is assigned to certain stems. If the distribution can be described with one conjugation class, we can easily motivate the extension of this pattern to new verbs. Ultimately, however, I demonstrate that such an approach is untenable.

Conceivably, the theme vowel could be added to the stem in a block as a rule of exponence, rather than directly as part of a stem. Since the theme vowels differ for each conjugation class, a rule of exponence involving theme vowels must be able to reference conjugation class. In PFM2 (Stump, 2015), properties like conjugation class must be mapped on to the property set using the **pm** function (property mapping), which is applied to stems to produce the correct form cell. To account for the distribution of the theme vowel (disregarding the third plural form with *-u-*), we might posit the following two rules of exponence in the first block (32):

(32) **Block 1**

- a. $X, \sigma : \{\text{pres ind conj } 3\} \rightarrow Xe$
- b. $X, \sigma : [\{\text{pres ind conj } Z\} \wedge [\{2\} \vee \{3 \text{ sg}\}]] \rightarrow X$

Here, we can see the regular theme vowel for third conjugation, *-e-* (see *gerere* ‘to bear’) expressed through the rule of exponence in (32-a). This rule will only apply to verbs to which the property *conj 3* has been mapped. The second rule in 3.26 produces

‘athematic’ forms as the stem is simply expressed as the output. It applies to stems that have had the property *conj* *Z* mapped on, where *Z* is the arbitrary marker of the ‘athematic’ conjugation.

The issue is made clear when considering how the mapping of the conjugation class properties applies to the individual property sets. **Stump** (2015: 115-119) describes how inflection class is assigned to the form cells via the **pm** function using Latin conjugation class as an example. The **pm** function applies to property sets to produce modified property sets. **Stump** (2015: 117) illustrates the result of the function as below in (33).

$$(33) \quad \mathbf{pm}_{\text{conj1}}(\{1\text{sg fut ipfv ind act}\}) = \{1\text{sg fut ipfv ind act } \mathbf{conj1}\} \quad (\text{Stump 2015: 117})$$

(33) demonstrates how the property set for the cell expressed by the 1SG FUTURE IMPERFECTIVE INDICATIVE ACTIVE has the property CONJ1 appended by means of the function $\mathbf{pm}_{\text{conj1}}$.⁸⁶ Stump generalizes this property mapping for all the different conjugation classes by defining the **Corr** function for Latin verbs:

$$(34) \quad \text{Property mapping } \mathbf{pm}_c \text{ for Latin verbs}$$

For any inflection-class index *c* and any well-formed morphosyntactic property set σ , $\mathbf{pm}_c(\sigma) = \sigma \cup \{c\}$ (Stump 2015: 118)

The definition in (34) simply formalizes the addition of any conjugation class feature to the property set. The aim of defining the conjugation class property *c* broadly is to formalize the generalization that all Latin stems adhere to certain conjugation classes. The framing in (34) in terms of a general variable *c* does exactly this.

⁸⁶This property mapping is ultimately responsible for the mapping a first conjugation future to a form. Thus, the first conjugation verb *parō* has a future with *-āb*: *parāb*-.

This definition of the Latin **pm_c** function is used in the definition of the Latin **Corr** function, seen below in (35):

$$(35) \quad \text{Where } \mathbf{Stem}(\langle L, \sigma \rangle) \text{ belongs to conjugation } c, \mathbf{Corr}(\langle L, \sigma \rangle) = \langle \mathbf{Stem}(\langle L, \sigma \rangle), \mathbf{pm}_c(\sigma) \rangle$$

The statement in (35) states that the **Corr** function, when applied to a specific lexeme (*L*) with a specified morphosyntactic property set (σ), produces the form correspondent of a content cell where the property mapping function has applied the property *c* to the property set σ .

If the lexeme FERRE only has one stem in the present indicative, represented below in (36), only one conjugation class can be assigned to the forms that use that stem.

$$(36) \quad \mathbf{Stem}(\langle \text{FERRE}, \sigma : \{\text{imperfective}\} \rangle) = \textit{fer}$$

To illustrate this, consider the derivations for the form cells of the 1PL *ferimus* (i.e. a thematic form) and the 2SG *fers* (i.e. an athematic form), represented in (37).

$$(37) \quad \begin{aligned} \text{a. } & \mathbf{Corr}(\langle \text{FERRE}, \{\text{imperfective pres ind act 1pl}\} \rangle) = \langle \mathbf{Stem}(\langle \text{FERRE}, \{\text{imperfective pres ind act 1pl}\} \rangle), \mathbf{pm}_c(\langle \text{FERRE}, \{\text{imperfective pres ind act 1pl conj3}\} \rangle) \rangle \\ \text{b. } & \mathbf{Corr}(\langle \text{FERRE}, \{\text{imperfective pres ind act 2sg}\} \rangle) = \langle \mathbf{Stem}(\langle \text{FERRE}, \{\text{imperfective pres ind act 2sg}\} \rangle), \mathbf{pm}_c(\langle \text{FERRE}, \{\text{imperfective pres ind act 2sg conj3}\} \rangle) \rangle \end{aligned}$$

There are two crucial observations to be gleaned from (37). Firstly, both the 1PL and 2SG correspond to the same stem, namely *fer*. The second observation is that the

property *conj3*, as per the **pm_c** function, is added to both the 1PL and 2SG cells. This addition is caused by the fact that they both use the same stem, in accordance with the **Corr** function defined in (35). Essentially, the **pm_c** function is inextricably linked to the stem choice.

This latter fact means that the two forms will not differ in terms of the application of the rules in (32), repeated in (38).

(38) **Block 1**

- a. $X, \sigma : \{\text{pres ind conj } 3\} \rightarrow Xe$
- b. $X, \sigma : [\{\text{pres ind conj } Z\} \wedge [\{2\} \vee \{3 \text{ sg}\}]] \rightarrow X$

The rule in (38-a) will apply to both thematic form 1PL *ferimus* and the athematic form 2SG *fers*, as they have the property [conj3] in the property sets.⁸⁷ Thus, after Block 1, the realization of both forms resembles the following:

- (39) a. Intended realization of Block 1 for 1PL *ferimus*: *fere*
 Actual: *fer*, $\sigma : \{\text{imperfective pres ind act 1pl}\} \rightarrow fere$
- b. Intended realization of Block 1 for 2SG *fers*: *fer*
 Actual: *fer*, $\sigma : \{\text{imperfective pres ind act 2sg}\} \rightarrow fere$

As we can see, the property mapping has caused the 2SG to be thematic, rather than athematic (*fers*).

To put simply, the mechanisms by which conjugation class is assigned to a stem are

⁸⁷One might imagine the possibility that the environment in (38-a) could be specified to include the first person and third plural. This would mean that the theme vowel would not be added to the second person and third singular. However, doing so would ultimately erode the generalization that *ferō* behaves as a third conjugation for the most part. We would need an entirely separate rule to apply to regular third conjugation rules, as those do add the theme vowel to the second person and third singular.

too limited to account for the distribution of the thematic vowel in *ferō*. Effectively, the stem *fer* can only receive one conjugation class property. Thus, PFM can only handle athematic-thematic alternations via stem functions which are dedicated to expressing the individual stems.

3.4.5.1 DISCUSSION OF PFM2 APPROACH

In summary, PFM2 cannot account for the pattern of athematicity in a unified way. As discussed in Section 3.4.1, the morphomic distribution of athematicity is reliant on detaching the property of being athematic from the exponent. However, PFM2 is unable to handle such an abstraction. In order to account for the distribution, each stem must be listed according to the environment it appears in, meaning the pattern of athematicity is purely accidental. In other words, there are no thematic or athematic stems, as they are all listed individually.⁸⁸

This is not to say that PFM2 cannot mechanically derive the individual cells of these verbs. As shown above, it is possible to list out the stems in such a way that they map onto the attested forms. However, it is impossible to generalize across all of the irregular verbs, as the athematicity does not coincide directly with the individual stems. To put it simply, the generalization proposed in Section 3.4.1 is too morphologically abstract for PFM2 to handle.

3.4.6 COMPARISON BETWEEN DM AND PFM2

Section 3.4.3 and Section 3.4.4 have described how DM and PFM2 respectively account for the athematic morpheme in Latin. With regards to the most crucial aspect of this

⁸⁸As an example, the stems *vī* and *vol* in the present indicative of *velle* can both be described as athematic, as the endings are simply attached to the stem without a theme vowel. However, the other stem in the present indicative, *volu*, is also analyzable as athematic in this criterion. The fact that it ends in a vowel is simply a coincidence. Thus the endings are also attached directly to the stem.

comparison, namely whether the frameworks encode the athematic generalization in a cohesive way, DM clearly demonstrates an advantage over PFM2. In DM, the cells which are athematic are directly referenced via an impoverishment rule. The idiosyncrasies of the individual verbs are expressed via their respective vocabulary items. In PFM2, there is no cohesive rule that can account for the generalization. Rather, the stem functions directly encode the stems, without reference to (a)thematicity. Attempts to describe the distribution in terms of conjugation class by decomposing the theme vowel prove inadequate in describing the data.

This specific aspect of the comparison is crucial, as encoding the generalization in a unified way motivates the extension of the pattern to new verbs. If the individual forms of the different verbs are described via specified stem functions, there is no way for the pattern to be extended. The DM analysis allows for this extension, as it provides a cohesive rule describing the distribution.⁸⁹

We may now return to the other questions posed in Section 3.4.2, restated below:

1. Do morphemes constitute evidence for the status of a distinct morphological component of grammar?
2. Is there a meaningful difference between how morpheme-based and stem-based approaches account for morphomic distributions?

In order to assess the first question, it is important to clarify what we would expect from each framework. Since DM does not make any explicit reference to a morphomic template, we would expect a DM-style approach to be unable to account for the athematic distribution as it patterns morphomically. PFM2, on the other hand, can encode morphomic properties directly via property mapping. Thus, we would expect the

⁸⁹For a discussion of how the extension to *nōlō* can be described in DM, see Appendix A.6.

framework to be able to describe the distribution via a unified property expressed in the property set.

The two analyses in Section 3.4.3 and Section 3.4.4 directly contradict the aforementioned expectation. The DM analysis is able to account for the distribution due to its ability to separate the property of athematicity from the exponent of the root. The PFM2 analysis demonstrates a crucial limitation, namely that morphomic distributions can only be encoded if they align perfectly with a stem. In the PFM2 analysis in Section 3.4.4, athematicity is simply a coincidence.

With regards to the second question, namely whether there is a difference between morphomic and stem-based approaches when it comes to accounting for morphomic distributions, we must look no further than the limitation in PFM2 that was just mentioned. A stem-based analysis explicitly requires a morphomic distribution to be expressed in terms of stems. Since athematicity is a morphological property which is distinct from stem-hood, it cannot be expressed via a stem-based framework like PFM2.

DM, on the other hand, uses a combination of impoverishment and Vocabulary Insertion to account for the morphomic distribution. The fact that the framework uses two distinct mechanics allows for the separation of athematicity as a property from the exponent of the root (or stem). This is a direct consequence of the fact that DISTRIBUTED MORPHOLOGY decomposes words fully into distinct morphemes. Thus, the analysis in Section 3.4.3 demonstrates the benefits of a piece-based framework when it comes to encoding morphomic patterns.

To summarize, I have argued that the DM account in Section 3.4.3 is equipped to handle the athematic morpheme, while the PFM2 account in Section 3.4.4 fails to account for the athematicity in a unified manner. As the morpheme is expressed primarily through the presence and absence of the theme vowel, a morpheme-based approach which can

explicitly reference these segments, as the DM account does with its impoverishment rule, can better account for the data than a stem-based approach.

3.5 CONCLUSION

This chapter investigated the morphomic distribution in the irregular verbs in Latin. The first part of the chapter was dedicated to discussing the diachronic origin of periphrasis in the Latin verbs *nōlō* and *mālō*, providing novel explanations for both the synchronic distribution of the forms in Plautus as well as the diachronic process by which the periphrasis was derived.

The subsequent part of the chapter discussed the synchronic distribution of athematicity, as demonstrated by several irregular verbs in Latin. This morphomic pattern poses a serious challenge for synchronic frameworks, as the pattern of athematicity, expressed in terms of an affixal alternation involving the theme vowel of the third conjugation, is highly idiosyncratic. I have demonstrated that a piece-based framework such as DISTRIBUTED MORPHOLOGY can describe the distribution, despite the expectation that the morphomic pattern can only be described in pure morphological terms. Additionally, the stem-based framework, PARADIGM FUNCTION MORPHOLOGY 2 is unable to account for the distribution, indicating that a dedicated morphological component of grammar is not required to describe morphomic patterns.

The next chapter focuses on how another morphomic distribution, the N-PATTERN, intersects with analogical change.

capèssèt? - a Bergamasque mobster (probably)

4

Irregular Distributions and Leveling

4.1 INTRODUCTION

As we have seen in Chapter 3, non-root intra-paradigmatic alternations have the potential to inform our understanding of diachronic change and synchronic theory. This chapter discusses another aspect which is relevant to irregular distributions of non-root intra-paradigmatic alternations, namely the analogical process of leveling. Given that these alternations necessarily involve two or more alternants, changes involving the replacement of one of the alternants with another is a common phenomenon, such that the

number of alternants in a paradigm is reduced. As with other examples of leveling, we can say uncontroversially that there is a bias against intra-paradigmatic alternations, causing the number of alternants to reduce over time.¹ What is less clear, however, is exactly what influences this pressure, and whether there is any directionality to these changes.

This chapter investigates how an irregular distribution in Romance, the N-PATTERN, is affected by leveling, with specific attention given to the affixal alternation involving the reflex of the Latin suffix *-sc-*. While the different attestations of this distribution are well-studied, particularly from a morphomic perspective (Maiden, 2003, 2018), how leveling has affected the distribution is a diachronic change which has not garnered as much attention. This is not surprising, as much of the focus of the study of morphemes has been to show their persistence over time (cf. Maiden 2018). However, as this chapter will show, it is clear that morphomic distributions can and indeed are affected by analogical change. A deeper inquiry into how leveling affects these distributions can thus provide new insights into diachronic change.²

4.1.1 N-PATTERN AND ITS DERIVATIVES

Before discussing the specific questions that this chapter aims to address, it is necessary to briefly describe the changes that affect the N-PATTERN. Later sections will discuss each stage in depth.

The N-PATTERN is a morphomic distribution affecting the present tense of many varieties of Romance. Table 4.1 demonstrates the parts of the paradigm for the Italian

¹Especially where the alternation is not phonologically conditioned.

²Throughout the chapter, I often use the phrase ‘leveling’ to refer to the elimination of an alternant in a given cell. It should be noted that an intra-paradigmatic alternation can be entirely leveled over time in favor of one of the alternants such that the alternation is eliminated altogether. However, even if an alternation is still present elsewhere within the paradigm, I may refer to a given cell as ‘leveled’.

verb *finire* ‘to finish’ which exhibits an alternation involving the suffix *-isc-*:³

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisco</i>	<i>finisci</i>	<i>finisce</i>	<i>finiamo</i>	<i>finite</i>	<i>finiscono</i>
PRES SUBJ	<i>finisca</i>	<i>finisca</i>	<i>finisca</i>	<i>finiamo</i>	<i>finiate</i>	<i>finiscano</i>
IMP	–	<i>finisci</i>	<i>finisca</i>	<i>finiamo</i>	<i>finite</i>	<i>finiscano</i>

Table 4.1: PRES IND/SUBJ of *finire* ‘to finish’ in Italian

Crucially, the suffix *-isc-* is found in the singular and the 3PL and is absent from the 1PL and 2PL. This distribution is the N-PATTERN.

The naming of this pattern can be attributed to Maiden and is used throughout his work on morphemes in Romance (Maiden, 2003, 2004, 2005, 2018); it has been subsequently adopted by authors writing about the distribution. The name is derived from the fact that if the forms in the present are placed in rows, the distribution of the cells with the reflex of *-sc-* resembles the letter N in Morse code. Consider the shaded cells in the present indicative of several Romance languages, seen below in Table 4.2.

1SG	2SG	3SG	1PL	2PL	3PL
<i>finisco</i>	<i>finisci</i>	<i>finisce</i>	<i>finiamo</i>	<i>finite</i>	<i>finiscono</i>
<i>patéixo</i>	<i>patéixes</i>	<i>patéix</i>	<i>patím</i>	<i>patío</i>	<i>patéixen</i>
<i>păţesc</i>	<i>păţeşti</i>	<i>păţească</i>	<i>păţim</i>	<i>păţiţi</i>	<i>păţească</i>

Table 4.2: Italian *finire*, Catalan *patir*, Romanian *a păţi*

The letter N in Morse code – ·, consisting of a dash followed by a dot, mirroring the shaded cells in Table 4.2. The forms in the singular correspond to the dash and the 3PL corresponds to the dot. The naming of the pattern after the letter N represents the apparent arbitrariness of the pattern (Maiden 2003: 2). For the rest of this chapter, I will follow the practice set forth by Maiden and will refer to this specific pattern as the N-PATTERN.

Returning to its distribution, there are two pathways of directional change that affect the N-PATTERN. The starting template, the N-PATTERN, is exemplified by a variety of

³As we will see, the suffix is completely absent from the rest of the paradigm outside of the present tense.

Romance languages, including Italian, Catalan, Romanian, and the Lombard dialects of Mantua and Gavardo, among others. This distribution has been adapted into several other minor patterns, diverging into two distinct paths.

The first path (henceforth PATHWAY 1), represented by the dialects of Lombardy, began with the extension of *-sc-* to the 1PL forms, before introducing the suffix *-sc-* throughout the subjunctive. This development is summarized schematically below, with the highlighted cells showing the forms containing the reflex of *-sc-*.

(1) PATHWAY 1

a. Stage 1: Lombard (Mantuan)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

b. Stage 2: Lombard (Cremonese, Bergamasque, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

c. Stage 3: Lombard (Milanese, Bresciano, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

The second path (henceforth PATHWAY 2), represented with the diagram in (2), began with a complete leveling of the *-sc-*-less forms in the subjunctive. Piedmontese uniquely introduced *-sc-* to the 2PL of both the indicative and subjunctive as well.

(2) PATHWAY 2

a. Stage 1: Catalan

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

b. Stage 2: Ligurian, Old Occitan, etc.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

c. Stage 3: Piedmontese

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

Both of these pathways of development represent a tendency to gradually introduce *-sc-* to all cells in the present tense.⁴

These two pathways present two notable subgeneralizations:

1. Wherever *-sc-* is introduced to the indicative, it is also introduced to the subjunctive (but crucially not vice versa).
2. Additionally, the indicative is more likely to retain an alternation between *-sc-* and *-Ø-*.

⁴Crucially, the forms in the non-highlighted cells are still found in other tenses. In fact, the forms not marked with the morphology specific to the highlighted cells dominate the cells outside of the present tense. In other words, forms with *-sc-* are limited to the present tense, as are forms of ‘to go’ with the stem *va(d)-* (cf. Section 4.2.4). This indicates that speakers are analyzing these features as markers of the present tense. Thus, the *-Ø-* alternant is not being eliminated throughout the entire paradigm. Rather, its presence in the present tense specifically is being restricted over time.

Thus, we do not find the following distributions:

(3) UNATTESTED DISTRIBUTIONS

a. **DISTRIBUTION 1:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

b. **DISTRIBUTION 2:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

c. **DISTRIBUTION 3:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

4.1.2 ISSUES AT STAKE

The distributions outlined above pose several questions from both diachronic and synchronic perspectives surrounding the nature of the template set up by Maiden (2003, 2018, *et seq.*).⁵

4.1.2.1 DIACHRONY: LEVELING

From a diachronic perspective, the N-PATTERN template is disrupted in various ways, seemingly creating new templates in the process. These developments pose multiple

⁵See Section 2.4 for an expanded explanation of the template as proposed by Maiden.

questions relevant for diachronic change. For example, the changes in the N-PATTERN are complicated by the fact that different phenomena in the languages (stress patterns, distributions of *-sc-*, suppletion in the verb ‘to go’) do not always change in the same way; the different phenomena seem to be subject to individual changes. One might ask how the morphomic distributions break down in different phenomena. If all the phenomena are part of the same morphome, why do they not necessarily change in lockstep with one another?

Alternatively, we must ask what exactly motivates leveling in these paradigms. Some approaches to leveling heavily favor reference to specific proportional analogies as set up with other paradigms (Hill 2007, Garrett 2008, *a.o.*), while others allow the possibility of a paradigm-internal motivation for leveling (Kiparsky 1992), whether due to a universal preference or otherwise.⁶ The developments discussed in this chapter are relevant to this question. Does the leveling of the *-sc-*less alternant reference specific proportions, or can it be explained via respect to internal coherence?

However, most crucial for this discussion is the directionality of change. Why do we find the distributions described in (1) and (2), but not those in (3)? In this chapter, I propose that the leveling of morphological alternates is subject to a set of biases, which are reflected in the directionality of change. Firstly, usage can indirectly cause alternations to be leveled in favor of default alternants. As we will see, if usage affects the input for learners by reducing the frequency of a specific alternant, an alternation can be leveled, changing the shape of the morphomic distribution. Secondly, I argue that learners exhibit morphological biases by which they organize alternants. As we have seen, many non-root intra-paradigmatic alternations are organized along clear morphosyntactic environments.⁷ While irregular (and morphomic) distributions do exist, they are

⁶See also the discussion in Fertig (2013: 71-76) about the nature of leveling.

⁷See Section 2.3.2.3.

not exempt from the pressure to conform to a more morphologically regular pattern. In this sense, identifying what these biases are is crucial in accounting for the distribution of leveling.

4.1.2.2 SYNCHRONY: MORPHOMIC COMPOSITION

Synchronically, one of the key questions involving morphomic distributions revolves around how a morpheme is encoded. Elsewhere, I have suggested that these templates exist as a set of morphological rules that reference morphosyntactic features.⁸ The diachronic development of the N-PATTERN with specific reference to the distribution of the *-sc-* suffix implies a ‘narrowing’ of sorts when it comes to the anomalous cells in the present tense. From a synchronic perspective, such a narrowing is intriguing as it indicates that learners are also narrowing the rules they use to account for the distributions.

In this chapter, I provide an analysis of the different stages of development based on previous work on the N-PATTERN by Embick (2016), further supporting the idea that frameworks that do not explicitly allow for morphomic templates can still account for morphomic distributions.

4.1.3 ROADMAP

The layout of the rest of the chapter is as follows. Section 4.2 provides the necessary background on the suffix *-sc-*, as well as its distribution in the N-PATTERN. The final portion of this section is dedicated to describing a generalization proposed by Embick (2016) which is crucial for understanding the subsequent motivation for the leveling seen in PATHWAY 1 and PATHWAY 2. Section 4.3 proposes a novel set of biases by which

⁸See the last portion of Chapter 3.

intra-paradigmatic leveling can occur, involving multiple factors like usage, frequency, and morphological coherence. These biases are evidenced by the different developmental pathways. Section 4.4 applies these biases to cases in which *-sc-* is extended in terms of person and number, while Section 4.5 does likewise for cases in which the subjunctive was leveled in favor of *-sc-*. Section 4.6 concludes.

4.2 BACKGROUND

The following sections provide necessary background on the N-PATTERN, beginning with a discussion of the history of the Latin suffix *-sc-*.

4.2.1 LATIN *-SC-* SUFFIX

This suffix has a long history in various Indo-European languages and a wide distribution across the different branches.⁹ Although its phonetic shape can be uncontroversially reconstructed to the proto-language as **-skē/o-*, its exact function is disputed. Throughout the family, the suffix has a broad set of uses, including making iteratives, inchoatives, causatives, and even desideratives. Anatolian and Greek demonstrate an iterative or habitual use, shown by Hittite *walhiš^hkizzi* ‘beats repeatedly’ and Ionic Greek *φεύγεσ^hκον* *pheúges^hkon* ‘they would (habitually) flee’, respectively. Tocharian demonstrates the causative use, e.g. Tocharian B *stamäsk-* ‘to put’ cf. *stäm-* ‘stand’, as does Greek on occasion: *πιπίσ^hκω* *pipískō* ‘I cause to drink’. The suffix retains a desiderative

⁹Before discussing the background of *-sc-* and its development in the different Romance languages, it is necessary to make a brief terminological note. A large amount of literature on Romance is devoted to the distribution of this suffix, which is also known variably as an infix, interfix, or augment. The reason for this variability is that the affix is regarded as essentially having no function and is instead an empty segment, similar to theme vowels. In fact, some synchronic analyses of Romance languages assume that *-sc-* has been reanalyzed as a theme vowel (e.g. Goryczka 2022). I depart from this tradition and will continue to refer to it as a suffix, in order to unite the type of alternations seen in Romance with those in other languages. In this sense, I see ‘suffix’ as a (relatively) theory-neutral description of the segment in question.

use minimally in certain verbs pertaining to asking or wishing, as in Lat. *poscō* ‘I demand’, Ved. *pr̥cchāti* ‘(s)he asks’. Latin, as will be discussed in more depth shortly, demonstrates an inchoative use, encoding a sense of ‘becoming *x*’. While some scholars have variously suggested that the inchoative use was the original use (Kurylowicz, 1964, Sihler, 1995, *a.o.*), an alternative view holds the desiderative function as the original use, with the iterative function developing later (Jasanoff, 2003, Fortson, 2011, Jasanoff, 2019, *a.o.*).

In addition to the varied distribution of the suffix in different branches, several specific verbs can solidly be reconstructed with **-skē/o-* due to forming clear word equations in several languages, including Latin:

- Lat. *poscō* ‘I demand’, Ved. *pr̥cchāti* ‘asks’, OHG *for̥scōn* ‘to inquire’
- Lat. *gnōscō* ‘I know’, Old Persian *xšnāsāti* ‘recognizes (subj)’, Greek γινώσκω *gignōskō* ‘I know’
- Lat. *misceō* ‘I mix’, Gk. μίσγω *mísgō* ‘I mix’, Irish *mescaim* ‘to mix’, Old English *misćian* ‘to mix’

In Latin, the full distribution of *-sc-* beyond these archaic forms is complex. As mentioned, the suffix was generally used to make inchoatives, mostly from statives ending in *-ē*. These derivatives are in the third conjugation: *calēre* ‘to be hot’ → *calēscere* ‘to become hot’, *frīgēre* ‘to be cold’ → *frīgēscere* ‘to become cold’, *albēre* ‘to be white’ → *albēscere* ‘to become white’. In addition to these deverbal examples, the suffix *-ēsc-*, the combination of *sc* and a reanalyzed vowel, was added to nominal formations to make inchoatives: *ignis* ‘fire’ → *ignēscere* ‘to take fire’, *fortis* ‘strong, brave’ → *fortēscere* ‘to become brave’. The suffix *-sc-* was otherwise added sparingly to other declensions, although examples are clearly attested, e.g. deponents from verbs in *-iō* *proficiō* ‘I make

progress' → *proficīscor* 'I set out' and the backformed derivative *īrāscōr* from *īrātus* 'angry'.¹⁰

An important fact is that forms with the suffix were limited to the *infectum*, namely the non-perfect forms.¹¹ The suffix does make an appearance in three perfect stems: *poscō* 'I demand' (*poposci*), *compescō* 'I restrain' (*compescui*), and *misceō* 'I mix' (*miscui*). However, its presence in these forms is hardly surprising, as the segment /-sc-/ in these forms was etymological and likely no longer decomposable in Latin, given the long prehistory of the root + suffix combination and the inseparability from the root. In this sense, they should be seen as archaisms rather than genuine examples of -sc- verbs.

In all, the general distribution of -sc- was as follows: -sc- formed third conjugation verbs with inchoative function from (second conjugation) statives in long -ē- as well as from denominal sources. Examples of -sc- added to non-second conjugation forms exist, but were rare.

4.2.2 PROTO-ROMANCE DEVELOPMENT OF -SC-

The development of -sc- into the modern Romance languages is inseparable from shifts affecting the conjugation classes.¹² While this development is ultimately secondary to the main discussion of the distribution of -sc- in this chapter, this section provides context for understanding the current scope of the suffix in the modern languages.

To reiterate, in Latin, the distribution of -sc- was correlated with third conjugation verbs.¹³ However, in the modern languages the suffix is usually associated with the

¹⁰For more examples of the suffix being added to verbs of the first and fourth conjugation, rather than the second, see the discussion of such forms in Meul (2016: 54-55).

¹¹This is predictable given that the semantics of the suffix did not allow for the perfective aspect. Past tense inchoatives were formed with the imperfect tense.

¹²For background on the Latin verbal system, including the conjugation classes, see Appendix B.1.2.

¹³See Appendix B.1.2.

fourth conjugation, e.g. It. *finisco* ‘I finish’, accompanied with the preceding theme vowel *-i-* which characterizes the fourth conjugation. Exactly how this change occurred is controversial.

As described in Meul (2016: 63-68), there are two proposals for this shift in the literature. The first is that *-sc-* transferred from the third to fourth conjugation due to a conflation between the second conjugation (from which *-sc-* verbs were derived) and the fourth conjugation. This conflation itself was based on the development of a phonetic similarity of the personal endings. Due to regular sound change, many of the endings of the second and fourth conjugations would have been identical (Meyer-Lübke, 1894, Lausberg, 1962, *a.o.*). For example, the 1SG of the present indicative would have become PRom. **-jo* < *-eō*, *-iō*, as well as the entire present subjunctive: PRom. **-jam* < *-eam*, *-iam*, **-jas* < *-eaś*, *-iās*, etc.¹⁴ The consequence of the sound changes was that much of the difference between the second and fourth conjugations in the present tense was largely eroded, as the theme vowel merged. This hypothesis assumes that this conflation would have allowed for a reanalysis in which verbs with *-sc-*, which were generally formed to second conjugation verbs, e.g. Lat. *calēre* ‘to be hot’ → *calēscere* ‘to become hot’, would have instead been formed to fourth conjugation verbs.

Thus, verbs that were originally second declension could become fourth declension; for example, Latin *flōrēre* ‘to blossom’ alongside its derivative *flōrēscere* ‘to blossom’ became It. *fiorire* ‘id.’ with the *-sc-* suffix found in the present tense: It. *fiorisco* ‘I blossom’. Together with verbs that moved from the second to the fourth conjugation, the phonological similarity also allowed for the production of forms with *-sc-* that did not originally have the suffix, such as Italian *finiscō* from Lat. *finire*.

The second proposal, as laid out by Meul (2016: 64-65), is that the shift from

¹⁴See Meul (2016: 63). See also the discussion of the L-PATTERN in Section 4.5.3.

the second to fourth conjugation was based on semantic function. Many verbs in the third conjugation formed with *-sc-* were intransitives, with the meaning ‘become X’ (e.g. *calēscere* ‘to become hot’).¹⁵ This proposal contrasts these verbs with deadjectival verbs in the fourth conjugation with the sense ‘to make X’ (e.g. *ūnīre* ‘to unify’ from *ūnus* ‘one’). The semantic functions of the two types of verbs would have complemented each other, allowing for a reanalysis of *-sc-* as a suffix deriving intransitive inchoatives from the transitive type.

Meul (2016) supports this hypothesis by noting that many fourth conjugation verbs with *-sc-* retain both intransitive and transitive functions. This would suggest that the connection of *-sc-* with a specific intransitive function was lost, with *-sc-* becoming a marker of fourth conjugation verbs instead. Some of these verbs, however, would have kept both functions. For example, French *maigrir* (cf. 1SG *maigris*) can mean both ‘to make thin’ and ‘to become thin’.

As it is difficult to pinpoint which proposal more accurately accounts for the distribution in the modern languages, I do not commit to either proposal here. Both accounts require additional assumptions beyond those which have been outlined above, making an analysis of the diachronic shift worthy of a dedicated discussion, which cannot be done here. A further issue in the diachrony of the Romance verbal system is the fate of the second and third conjugations as a whole, which cannot be ignored in the context of this specific change. I will not, however, discuss that further here. For the purposes of this chapter, it is sufficient to be aware that this shift occurred, and accordingly that the vast majority of examples that contain the *-sc-* suffix in the following discussion are fourth conjugation, instead of the third conjugation.

¹⁵The verbs to which *-sc-* was added were often also intransitive (e.g. *calēre* ‘to be hot’).

IS THE SUFFIX -SC- OR -ISC-? As a terminological note, much of the literature discussing the suffix in Romance refers to it as *-isc-* or *-esc-* based on the fact that it is necessarily accompanied by a vowel (mostly *-i-*). In this chapter, I refer to the suffix as the *-sc-* suffix, but may refer to its outcomes together with the vowel. For example, the outcome of the suffix in Italian is *-isc-*: *finisco* ‘I finish’.

4.2.3 N-PATTERN

The Romance languages all inherited the suffix *-sc-* from Latin, although the distribution of the suffix differs depending on the language. What is crucial for our purposes is that some varieties show an intra-paradigmatic alternation in the present, as we have already seen in Section 2.3.2.4. This section discusses one of those distributions in depth, namely the N-PATTERN.¹⁶

The distribution of the N-PATTERN can be demonstrated in the following paradigms:

	ITALIAN ‘to finish’	CATALAN ‘to suffer’	ROMANIAN ‘to suffer’
1SG	<i>finisco</i>	<i>patéix</i>	<i>păţesc</i>
2SG	<i>finisci</i>	<i>patéixes</i>	<i>păţeşti</i>
3SG	<i>finisce</i>	<i>patéix</i>	<i>păţească</i>
1PL	<i>finiamo</i>	<i>patim</i>	<i>păţim</i>
2PL	<i>finite</i>	<i>patio</i>	<i>păţiţi</i>
3PL	<i>finiscono</i>	<i>patéixen</i>	<i>păţească</i>

Table 4.3: N-PATTERN in Italian, Catalan, and Romanian (Meul 2016: 81)

In these languages, the descendent of Latin *-sc-*, namely Italian *-isc-*, Catalan *-eix-*, and Romanian *-esc-/eşc-/easc-*, is found in all singular forms as well as the 3PL of certain fourth conjugation verbs. It is absent, however, from the 1PL and 2PL. The same distribution is also found in the present subjunctive and the present imperative in the corresponding cells; however, the suffix is otherwise absent from other tenses and

¹⁶For more discussion of distributions which do not involve an intra-paradigmatic alternation, see Appendix B.3.

moods.

This distribution differs remarkably from that found in Latin. The suffix is no longer found in other forms descended from the *infectum*; for example, the imperfect indicative which otherwise descends directly from the imperfect indicative in Latin lacks the suffix in these languages. Distributional gaps such as this imply that the suffix underwent a significant amount of paradigmatic restriction in terms of tense and mood. While the absence of the suffix from tenses like the imperfect can otherwise be explained,¹⁷ the absence of the suffix in the 1PL and 2PL of the present tense are particularly peculiar, given that the suffix was present in those environments in Latin.

The curious distribution found in these languages is made even more surprising given that this pattern is also found with other types of morphological and phonological expression. The N-PATTERN distribution is not just limited to the expression of the suffix *-sc-*, but can also be found in present tense formations of completely separate verbal constructions. Notably, the N-PATTERN is found with vowel alternations in the verb stem, other affixal alternations, and root suppletion alternations.

Among these, the most extensive type of alternations are the vowel alternations. Depending on the language, vowels in stressed syllables undergo various changes, including diphthongization, raising, and fronting (cf. Maiden 2018). Vowel changes of these types in the N-PATTERN distribution can be seen throughout Romance, including in languages that do not otherwise have the *-sc-* alternation following the N-PATTERN, as seen in Table 4.4.

Some of these paradigms involve alternations between diphthongs and monophthongs, as in Italian *-uo-* and *-o-*, while others involve a change in vowel quality, such as the alternation between *-i-* and *-e-* in Romanian.

¹⁷The suffix *-sc-* was likely reanalyzed as a marker of the present in these varieties. This explains its absence from the imperfect.

	ITALIAN 'to die'	ROMANIAN 'to come'	FRENCH 'to owe'	SPANISH 'to lose'
1SG	<i>muoio</i>	<i>vin</i>	<i>dois</i>	<i>pierdo</i>
2SG	<i>muori</i>	<i>vi</i>	<i>dois</i>	<i>pierdes</i>
3SG	<i>muore</i>	<i>vine</i>	<i>dois</i>	<i>pierde</i>
1PL	<i>moriemo</i>	<i>venim</i>	<i>devons</i>	<i>perdemos</i>
2PL	<i>morite</i>	<i>veniți</i>	<i>devez</i>	<i>perdeis</i>
3PL	<i>muiono</i>	<i>vin</i>	<i>doivent</i>	<i>pierden</i>

Table 4.4: Vowel Alternations in Stems in Romance Maiden (2018: 170-171)

The distribution of these vowel alternations is uncontroversially linked to stress (Meyer-Lübke, 1894, Meul, 2010, 2016, Maiden, 2018). In fact, the development of the vowel alternations in Table 4.4 neatly fall out from the assignment of stress in Classical Latin. As described in Weiss (2020: 110), stress in Classical Latin was assigned to the first syllable of a disyllabic word and for a longer word, on the penult if heavy or on the antepenult otherwise. For the different conjugations, this leads to the stress distribution for the present tense seen in Table 4.5, adapted from Maiden (2018: 168).

	1ST CONJ	2ND CONJ	3RD CONJ	4TH CONJ
INF	<i>portāre</i> 'bear'	<i>vidēre</i> 'see'	<i>pērdere</i> 'lose'	<i>dormīre</i> 'sleep'
	PRES IND			
1SG	<i>pórtō</i>	<i>videō</i>	<i>pērdō</i>	<i>dórmio</i>
2SG	<i>pórtās</i>	<i>vidēs</i>	<i>pērdīs</i>	<i>dórmīs</i>
3SG	<i>pórtat</i>	<i>videt</i>	<i>pērdit</i>	<i>dórmīt</i>
1PL	<i>portāmus</i>	<i>vidēmus</i>	<i>pērdimus</i>	<i>dormīmus</i>
2PL	<i>portātis</i>	<i>vidētis</i>	<i>pērditis</i>	<i>dormītis</i>
3PL	<i>pórtant</i>	<i>vident</i>	<i>pērdunt</i>	<i>dórmunt</i>
	PRES SUBJ			
1SG	<i>pórtem</i>	<i>videam</i>	<i>pērdam</i>	<i>dórmiam</i>
2SG	<i>pórtēs</i>	<i>videās</i>	<i>pērdās</i>	<i>dórmīās</i>
3SG	<i>pórtet</i>	<i>videat</i>	<i>pērdat</i>	<i>dórmiat</i>
1PL	<i>portēmus</i>	<i>videāmus</i>	<i>pērdāmus</i>	<i>dormiāmus</i>
2PL	<i>portētis</i>	<i>videātis</i>	<i>pērdātis</i>	<i>dormiātis</i>
3PL	<i>pórtent</i>	<i>videant</i>	<i>pērdant</i>	<i>dórmiant</i>

Table 4.5: Stress in Latin conjugation

Crucially, in the first, second, and fourth conjugations, the 1PL and 2PL have a long

vowel in the penult in both the indicative and subjunctive. According to Latin stress assignment, the stress falls on these long vowels, leading in turn to a stress pattern in these conjugations which directly conditioned the distribution of the vowel alternations in Table 4.4; the 1PL and 2PL are differentiated from the singular and 3PL forms in terms of whether the root is stressed. In the third conjugation, the present subjunctive has the same stress pattern. However, the indicative is stressed on the root throughout the paradigm, due to having a short vowel in the penult.

It should be noted, however, that in most branches of Romance apart from Romanian, the present indicative of third conjugation verbs adopts the same stress pattern as in the other conjugations (Maiden 2018: 168). This regularization is seen in Table 4.6, which shows the outcome of the third conjugation verb, Lat. *perdere* in several of the modern languages.

	1SG	2SG	3SG	1PL	2PL	3PL
LATIN	<i>pérdō</i>	<i>pérdis</i>	<i>pérdit</i>	<i>pérdimus</i>	<i>pérditis</i>	<i>pérdunt</i>
ITALIAN	<i>pérdo</i>	<i>pérdi</i>	<i>pérde</i>	<i>perdiámo</i>	<i>perdéte</i>	<i>pérdono</i>
SPANISH	<i>piérdo</i>	<i>piérdes</i>	<i>piérde</i>	<i>perdémos</i>	<i>perdís</i>	<i>piérden</i>
ROMANIAN	<i>piérd</i>	<i>piérzi</i>	<i>piérde</i>	<i>piérdem</i>	<i>piérdeți</i>	<i>piérd</i>

Table 4.6: ‘to lose’ in Latin, Italian, Spanish, and Romanian

Only Romanian preserves the original stress pattern, while Italian and Spanish have shifted the stress onto the ending in parallel with other conjugations. This shift has also affected the outcomes of the vowel (cf. Table 4.5). The key takeaway from these forms is that even in the cases where Latin had a stressed root throughout the present indicative, there was a tendency for the daughter languages to change the stress pattern in the third conjugation in favor of the pattern found in the other conjugations, namely the N-PATTERN.

These vowel alternations have also exerted analogical influence on other forms where one would not expect vowel changes. In Spanish, for example, roots with etymological -ē-

and $-\bar{o}$ - occasionally diphthongize unexpectedly due to analogy with forms that originally had short vowels.¹⁸ The Latin short vowels $-\check{e}$ - and $-\check{o}$ - were regularly diphthongized when stressed, producing the diphthongal alternation in an N-PATTERN: Lat. *rógō*, *rógās*, etc. ‘ask’ > Sp. *ruego*, *ruegas*, etc. but Lat. *rogāmus*, *rogātis* > Sp. *rogamos*, *rogáis*.¹⁹ The long vowels $-\bar{e}$ - and $-\bar{o}$ -, on the other hand, were not diphthongized when stressed.²⁰ Thus, for Latin *mōnstrō*, *mōnstrās*, etc. ‘show’, we would expect to find **mostro*, *mostras*, etc. in Spanish. Instead the present has been remodeled as if the vowel were short: *muestro*, *muestras*, *muestra*, *muestran* in the singular and 3PL and *mostramos*, *mostráis* in the 1PL and 2PL. Other languages like Portuguese and Romanian show similar analogies, although the exact details and analogical proportions are language specific (Maiden 2018: 174).

If the N-PATTERN was only expressed in terms of stress, one might be tempted to assume that the distribution is simply phonological. However, the following case suggests that the distribution is encoded morphologically in these languages. In addition to the alternations in vowel quality and diphthongization, the N-PATTERN is also exhibited by examples of root suppletion, specifically in the verb ‘to go’. Many of the daughter languages have distributed three roots, Lat. *vādere* ‘to go’, Lat. *īre* ‘to go’, and Lat. *ambulāre* ‘to traverse’, in various suppletive patterns that follow the same N-

¹⁸In what follows, I gloss over the fact that Vulgar Latin had already lost the length distinction. Instead, long and short vowels were distinguished in vowel quality. This new vowel system was itself subject to changes, e.g. the merger between Lat. $-\bar{e}$ - and $-\check{e}$ - into VL $-e$ -.

We can summarize the changes to the relevant Latin vowels in Spanish as follows: CL. $-\check{e}$ -, $-\check{o}$ -, $-\bar{e}$ -, $-\bar{o}$ - > VL $-\text{e}$ - [e], $-\text{o}$ - [o], $-e$ -, $-o$ - (Weiss, 2020). When stressed VL $-\text{e}$ - and $-\text{o}$ - become $-ie$ - and $-uo$ - respectively in Spanish, and $-e$ - and $-o$ - when unstressed (Penny, 2002).

As discussed above, the 1PL and 2PL in the third conjugation have had their stress patterns remodeled after other conjugations. In Table 4.6, this culminates in an unstressed root in *perdemos* and *perdēte*, where Lat. $-\check{e}$ has become $-e$ -, producing the alternation between $-ie$ - and $-e$ -.

¹⁹We are specifically interested in the vowel of the root in these forms. I therefore put aside any discussion of the endings.

²⁰Under very specific conditions, the outcome of a Latin $-\bar{o}$ - would diphthongize due to sound change, although this was dependent on a following palatal causing metaphony, e.g. Lat. *cicōnia* ‘stork’ > Sp. *cigüeña*.

PATTERN template illustrated above for paradigms containing the suffix *-sc-* as well as vowel alternations within the root.²¹ Some examples can be seen below in Table 4.7.

	Italian	Catalan	French	Old Spanish
1SG	<i>vàdo</i>	<i>vaig</i>	<i>vais</i>	<i>voy</i>
2SG	<i>vài</i>	<i>vas</i>	<i>vas</i>	<i>vas</i>
3SG	<i>và</i>	<i>va</i>	<i>va</i>	<i>va</i>
1PL	<i>andiàmo</i>	<i>anem</i>	<i>allons</i>	<i>imos</i>
2PL	<i>andàte</i>	<i>aneu</i>	<i>allez</i>	<i>ides</i>
3PL	<i>vànno</i>	<i>van</i>	<i>vont</i>	<i>van</i>

Table 4.7: Suppletive alternations for ‘to go’ in Romance (Maiden 2004: 232-233)

In each of these languages, the reflex of *vādere* is seen throughout the singular and 3PL. In Italian, Catalan, and French, the 1PL and 2PL show reflexes of *ambulō*, while Old Spanish has reflexes of *īre* in these cells.

Another interesting example from Italian is the suppletive alternation for the verb *uscire* ‘to go out’. Here, the stem *esc-* from Latin *exīre* ‘to go out’ alternates with the stem *usc-*, originally from Latin *ōstium* ‘door’, as seen in Table 4.8.²²

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>èsko</i>	<i>èsci</i>	<i>èsce</i>	<i>usciamo</i>	<i>uscite</i>	<i>èsono</i>
PRES SUBJ	<i>èsca</i>	<i>èsca</i>	<i>èsca</i>	<i>usciamo</i>	<i>usciate</i>	<i>èscano</i>

Table 4.8: PRES IND/SUBJ of *uscire* ‘to go out’ in Italian (Maiden 2018: 203)

Maiden (2018) details several other minor examples of verbal suppletion that affects roots beyond the verb for ‘to go’ in various Romance languages. However, for the rest of the chapter, I focus on the reflexes of the verb(s) for ‘to go’ as they are the most

²¹Maiden (2018) assumes that It. *andare* goes back to *ambulāre*, although this word has a notoriously difficult etymology.

²²[ks] became [ʃ] intervocalically before [i]; thus we find *esc-* instead of *ess-* (cf. *lusso* < Lat. *luxus*). Why the initial *e-* is still present in this form is more difficult (*esco* versus **sco*). Tekavčić (1980: 76), as cited in Maiden (2018: 203), suggests that the [u] is rounded due to the following [ʃ]; thus, *usc-* could be from Lat. *exīre*. See Maiden (2018: 202-203) for further discussion of the etymology of these forms. Regardless of the exact origin of the respective alternants, it is clear that there is a suppletive alternation in this paradigm that at least from a synchronic perspective is not governed by regular phonology.

consistent representatives of the N-PATTERN.

The final example of how the N-PATTERN has been extended is the distribution of another suffix *-idi-*. This suffix was originally borrowed into Latin from the Greek denominative verbal suffix *-ιζ-* *-iz-*: VL *bapitidiāre* < Gk. βαπτίζω *baptízō* ‘to baptize’. The Greek suffix *-ιζ-* was derived historically from a combination of nominal stems ending in *-ιδ-* *-id-* or *-ιγ-* *-ig-* and the PIE present stem suffix *-ie/o-*. For example, the denominative verb ἐλπίζω *elpízō* ‘to hope’ was derived from the noun ἐλπίς *elpís* ‘hope’ (ἐλπίδος *elpídos* GEN SG).

In Latin, the suffix *-ιζ-* was originally borrowed as *-iss-* in early loans: *badissāre* < Gk. βαδίζω *badízō* ‘to walk’. The spelling eventually changed to *-z-*, as in *baptizāre* < Gk. βαπτίζω *baptízō* ‘to baptize’, before becoming interchangeable with spellings in *-idi-*. Some argue that the spellings of *-z-* and *-di-* represent either [dz] or [dʒ], as both were used for the Greek letter *-ζ-*: *oridia/oryza* < Gk. ὀρυζᾶ *óryza* ‘rice’ (cf. Herman 2000, Alkire & Rosen 2010, *a.o.*). Verbs with this suffix were situated in the first conjugation in Latin, as noted by the above examples. The daughter languages for the most part generalized the suffix *-idi-* as what Meul (2016) calls a ‘lexically-integrated morpheme’; in other words, the suffix came to mark the verbal class of certain lexical items, which are all in the first conjugation. Among the languages which distribute this suffix in the N-PATTERN are Romanian, Istrian, Dalmatian, some Northern Italian dialects, varieties of Ladin and Friulian, as well as some Southern Italian languages like Corsican (Meul, 2016, Maiden, 2018).

Interestingly, some of the languages which otherwise have a strong affinity for the N-PATTERN do not distribute the suffix *-idi-* in this way. Compare the reflexes in Rom. *-ez-* with those in It. *-eggi-(are)*, Sp. *-e-(ar)*, Fr. *-oy-(er)*, and Port. *-e(j)-(ar)* (Meul 2016: 72), which have come to show the reflex of *-idi-* throughout the entire paradigm:

	ROMANIAN	ITALIAN	FRENCH	SPANISH	PORTUGUESE
	‘to work’	‘to ripple’	‘to feast’	‘to cry’	‘to flame’
1SG	<i>lucrez</i>	<i>ondeggio</i>	<i>festoie</i>	<i>lagrimeo</i>	<i>chamejo</i>
2SG	<i>lucrezi</i>	<i>ondeggi</i>	<i>festoies</i>	<i>lagrimeas</i>	<i>chamejas</i>
3SG	<i>lucrează</i>	<i>ondeggia</i>	<i>festoie</i>	<i>lagrimea</i>	<i>chameja</i>
1PL	<i>lucrăm</i>	<i>ondeggiamo</i>	<i>festoyons</i>	<i>lagrimeamos</i>	<i>chamejamos</i>
2PL	<i>lucrați</i>	<i>ondeggiate</i>	<i>festoyez</i>	<i>lagrimeáis</i>	<i>chamejais</i>
3PL	<i>lucrează</i>	<i>ondeggiano</i>	<i>festoient</i>	<i>lagrimean</i>	<i>chamejam</i>

Table 4.9: Reflexes of *-idi-* in Romance

In summary, the N-PATTERN is a distribution of various morphological expression that is based on an alternation in the present tense, where the 1PL and 2PL alternate with the other cells of the paradigm. The diverse types of morphological expression involved, including stress, suppletion, phonologically-induced stem changes, and affixation, indicate the existence of some abstraction which can account for why all these dissimilar phenomena have a unified distribution.

THE SUFFIX *-ISC-* AND THE N-PATTERN As we have seen, several Romance languages (e.g. Italian) distribute the reflex of the *-sc-* suffix into the N-PATTERN in the present indicative and subjunctive. We have also seen that the original distribution of *-sc-* did not have this distribution. It was limited to the *infectum*, which included the present indicative and subjunctive; however, it was present in the first and second plural, as well as several additional tenses (e.g. future, imperfect). In this sense, it is clear that those languages which have distributed *-sc-* into the N-PATTERN have greatly restricted the distribution of the suffix. This situation has led to a question that has generated much discussion in the literature: what caused this restriction?

It is clear from the outset that those cases where *-sc-* is found only in the singular and 3PL must be tied with other examples of the N-PATTERN. There are two main approaches in the literature. There are two types of approach in the literature: phonological and morphological. Ultimately, I adopt the general diachronic development proposed by Maiden (2003), which involves the extension of the N-PATTERN to verbs with *-sc-* on the basis of a morphological comparison. Examples of the phonological approach (e.g. DiFabio 1990, Anderson 2008) propose that the extension of the N-PATTERN to verbs with *-sc-* is due to a phonological restriction involving underlying stress. However, this approach has been widely disputed (cf. Meul 2010, 2016, Maiden 2018, *a.o.*).²³

The basic principle of Maiden (2003) is that once *-sc-* was introduced to the fourth conjugation, its distribution was restrained in many of the languages on the basis of the N-PATTERN. In his works, Maiden refers to the concept of the *morphome* as what is responsible for the extension.²⁴ Essentially, the N-PATTERN was abstracted from cases in which it was formerly phonologically conditioned (see above), before being extended. For verbs with *-sc-*, an alternation was created based on the presence and absence of the suffix which fit this abstraction.

This is largely the position that I will adopt for the rest of the chapter. However, as we will see, I do depart from Maiden in terms of the nature of this abstraction. The next section expands on the notion of the N-PATTERN as a template.

4.2.3.1 N-PATTERN AS A MORPHOMIC TEMPLATE

Further discussion of the N-PATTERN hinges upon the question of how the distribution of *-sc-* as well as the other phenomena, namely vowel alternations, root suppletion,

²³I refer the reader to these references for detailed criticisms of the phonological approach.

²⁴See Section 2.4 for more information on morphomes.

and the suffix *-idi-*, are related to each other, both from a diachronic and synchronic perspective. In other words, how are the different phenomena that have N-PATTERN distributions related to one another?

The facts outlined above in Section 4.2.3 have led Maiden (2003, 2004, 2005, 2017, 2018, *et seq.*) to propose that Romance languages follow a morphological template, known as a morphome. Here, the N-PATTERN represents a morphome where certain arbitrary cells are mapped together synchronically, specifically containing the forms of the present singular and the 3PL. For Maiden, who argues for the synchronic existence of morphomes, speakers of Romance languages have organized the intra-paradigmatic alternations like changes in root vowels (Table 4.4), root suppletion (Table 4.7), and suffixal alternations like *-idi-* and *-sc-* along a specific shape, the N-PATTERN, within their mental grammar.

Although Maiden does not commit to any specific formalism regarding the nature of the morphome, it can be conceived of as a simple morphological abstraction that exists in the minds of speakers. However, Maiden does not focus on the synchronic evidence for such mental structures. Instead, the main justification for positing that morphomes are psychologically real morphological objects comes in the form of diachronic, rather than synchronic evidence (Maiden 2018: 10). As a result, the diachronic changes that affect morphomic distributions are key to the discussion of whether morphomes constitute real mental structures.²⁵

Changes that center around morphomes play a crucial role in Maiden's analysis. We can exemplify the importance of these changes through the application of the N-PATTERN to novel phenomena in Romance. In diachronic terms, Maiden (2018) proposes the following development in Romance:

²⁵See the discussion in Section 2.4.

1. Firstly, the phonologically-conditioned stress pattern was morphologized. This effectively distinguished the 1PL and 2PL from the other persons in the present indicative and subjunctive in many of the conjugations. Phonological changes that affected vowel quality or caused diphthongization, as seen in Table 4.4, directly coincided with this morphologization. These stem alternations for Maiden served as the main evidence for speakers to distinguish the 1PL and 2PL from the other person and number combinations.
2. Speakers then abstracted the distribution of the phonologically-induced stem alternations as the N-PATTERN.
3. This pattern was then extended, either in Proto-Romance or soon thereafter, to other paradigms, including those involving the suffixes *-sc-* and *-idi-* and the suppletive paradigm of the verb ‘to go’.
4. The surface variation between different Romance languages depended on how far speakers extended their mental representations of the morpheme to new phenomena. In the individual daughter languages, the morpheme itself was variously affected, with some, like Italian, preserving the original distribution, while others strongly limited its influence, e.g. French.

The application of the N-PATTERN in Romance to different morphological phenomena supports the existence of some sort of synchronic mental representation that encodes the distributional pattern. There are limits, however, to this approach. The diachronic evidence predominately demonstrates that speakers *can* make distributional abstractions and apply these to new phenomena, but does not specifically tell us how these distributions are encoded synchronically as objects in the mental representation in speakers’ minds. Crucially, there is room to see whether the diachronic changes

that affect morphomic distributions do in fact say something deeper about the mental encoding than the descriptive analysis proposed by Maiden. This is precisely what this chapter examines, with a specific emphasis on how leveling affects the irregular distribution of morphemes.

4.2.4 SYNCHRONIC ASSUMPTIONS: EMBICK 2016

Before discussing the main proposal for how biases affect analogical leveling, it is necessary to discuss the synchronic framework that I am assuming, namely an analysis of the N-PATTERN in Italian by Embick (2016). The reason for this need is twofold. First, Embick makes an observation about the distribution of alternants in the N-PATTERN which plays role in the upcoming diachronic analysis.²⁶ Second, the analysis provides a formalism which can be used to describe the different stages of leveling.

Embick (2016) proposes an analysis for the distribution of the N-PATTERN in Italian based on the paradigms for *finisco* ‘I finish’ and *vado* ‘I go’ without referring to a morphomic template like that proposed by Maiden (2003, 2005, 2018, *et seq.*). Before discussing the details of the analysis, there is an important observation to be made about the structure of the verbal system in Romance. Despite the anomalous distribution of the N-PATTERN, the alternations in the present tense obscure a bigger generalization in which the present tenses are differentiated from the non-present tenses by overt morphological marking.

GENERALIZATION: Embick’s analysis focuses on the generalization that the exponents of the 1PL and 2PL in the present tense are consistent with the morphology found in other tenses. The exponence in the singular and 3PL, however, are unique to the present

²⁶As we will see, this observation is useful beyond accounting for the distribution from a synchronic perspective.

tense (besides the 1PL and 2PL).

To reiterate the facts from Section 4.2.3, the distribution of the forms with *-isc-* as well as the suppletive alternation of *va(d)-/and-* in the irregular verb *andare* ‘to go’ follow the N-PATTERN. I repeat the distribution below (adapted from Embick 2016: 290):

	FINIRE		ANDARE	
	PRES IND	PRES SUBJ	PRES IND	PRES SUBJ
1SG	<i>finisco</i>	<i>finisca</i>	<i>vado</i>	<i>vada</i>
2SG	<i>finisci</i>	<i>finisca</i>	<i>vai</i>	<i>vada</i>
3SG	<i>finisce</i>	<i>finisca</i>	<i>và</i>	<i>vada</i>
1PL	<i>finiamo</i>	<i>finiamo</i>	<i>andiamo</i>	<i>andiamo</i>
2PL	<i>finite</i>	<i>finiate</i>	<i>andate</i>	<i>andiate</i>
3PL	<i>finiscono</i>	<i>finiscano</i>	<i>vanno</i>	<i>vadano</i>

Table 4.10: N-PATTERN in Italian

The noteworthy part of the distribution in Table 4.10 is that the stems throughout the singular and the 3PL are aligned (*finisc-* and *va(d)-*) in opposition with the stems in the 1PL and 2PL (*fin-* and *and-*).

When presented with just the present tense, as above, one may get the impression that it is the morphological marking of the 1PL and 2PL that is anomalous. However, this is not the case when examining the full paradigms of these verbs. To anticipate the analysis outlined below, Embick (2016) builds on the observation that the forms in the 1PL and 2PL, despite appearing in the minority of cells in the present indicative and subjunctive, actually are seen throughout the majority of other tenses and moods. To exemplify this, consider the forms of *andare* and *finire* in Table 4.11 and Table 4.12 respectively.

	PRES IND	PRES SUBJ	IMPF	PRET	FUT
1SG	vàdo	vàda	<i>andàvo</i>	<i>andàì</i>	<i>andrò</i>
2SG	vài	vàda	<i>andàvi</i>	<i>andàsti</i>	<i>andràì</i>
3SG	và	vàda	<i>andàva</i>	<i>andò</i>	<i>andrà</i>
1PL	<i>andiàmo</i>	<i>andiàmo</i>	<i>andavàmo</i>	<i>andàmmo</i>	<i>andrémo</i>
2PL	<i>andàte</i>	<i>andiàte</i>	<i>andavàte</i>	<i>andàste</i>	<i>andréte</i>
3PL	vànnò	vàdano	<i>andàvano</i>	<i>andàrono</i>	<i>andrànnò</i>

Table 4.11: Forms of ANDÀRE ‘to go’ in Italian

	PRES IND	PRES SUBJ	IMPF	PRET	FUT
1SG	finisco	finisca	<i>finivo</i>	<i>finìi</i>	<i>finirò</i>
2SG	finisci	finisca	<i>finivi</i>	<i>finisti</i>	<i>finiràì</i>
3SG	finisce	finisca	<i>finiva</i>	<i>finì</i>	<i>finirà</i>
1PL	<i>finiàmo</i>	<i>finiàmo</i>	<i>finivàmo</i>	<i>finìmmo</i>	<i>finirémo</i>
2PL	<i>finìte</i>	<i>finiàte</i>	<i>finivàte</i>	<i>finìste</i>	<i>finiréte</i>
3PL	finiscono	finiscano	<i>finivano</i>	<i>finirono</i>	<i>finirànnò</i>

Table 4.12: Forms of FINIRE ‘to finish’ in Italian

The distribution of the stems in these paradigms indicates two key points. Firstly, the 1PL and 2PL of the present indicative and present subjunctive are more representative of the regular stem distribution. This plays a significant role in the synchronic analysis of Embick (2016), in which the stem found outside of the present tense is the elsewhere case.

Secondly, the marking in the rest of the present tense, exemplified by the bolded forms with *-sc-* and the stem *va(d)-*, serves as the main overt distinction between the present and non-present tenses. In other words, that the morphology is restricted to the finite present system signals the morphosyntactic features of the present to a listener.

To put it simply, for *andàre*, the root that expresses the present tense is *vad-*, while the *and-* stem is used elsewhere; likewise for *finìre*, the presence of *-isc-* is conditioned by the fact that the forms are in the present, while the forms with *-i-* are found elsewhere. If we put aside the forms of the 1PL and 2PL, this generalization is clear and is relatively uncontroversial among synchronic morphological frameworks.

ANALYSIS: Moving on to the analysis, Embick (2016) formalizes the intuition that the non-present forms are found in the elsewhere case (i.e. the default) in the framework of Distributed Morphology (DM).²⁷ For *andàre*, the *and* stem is the elsewhere case while the *va(d)* stem is specified for the present tense, which is expressed via a [-past] feature on the tense node. For the purposes of clarity, I have replaced the [-past] node with a [+pres] feature.²⁸ The two Vocabulary Items responsible are outlined below:

- (4) Vocabulary Items relevant for *andàre* (Altered from Embick 2016: 291)
- a. $v_{go} \leftrightarrow va(d) / \text{ ___ } [+pres]$
 - b. $v_{go} \leftrightarrow and$

A parallel set of rules express the distribution of the *-isc-* morpheme in *finìre*. Here, the present tense conditions the presence of *-isc-*, similar to the *vad* stem of *andàre*. The forms with *-i-* on the other hand, represent the elsewhere case. Thus, (5) demonstrates the VI rules relevant for the *-isc-* suffix, expressed as a verbalizer specified by a diacritic for the third conjugation.

²⁷See Section 3.4.3.

²⁸In effect, we can express the same distribution by replacing the T[-past] node with a [+pres] feature that is required to get *va(d)-* on the surface. The justifications for using T[-past] are syntactic in nature and are largely not necessary for our purposes. For this reason, I do not expand on the underlying structure assumed by Embick (2016). See the referenced work for more specific details.

(5) Vocabulary Items relevant for *-isc-* (Altered from Embick 2016: 291)

a. $v[\text{III}] \leftrightarrow \text{-isc-} / \text{ ______ } [+pres]$

b. $v[\text{III}] \leftrightarrow \text{-i}$

The benefit of positing these VI rules is that they directly account for the generalization that the *va(d)* stem and the *-isc-* suffix only occur in the present tense.²⁹

The main issue at hand of course is how we account for the fact that the supposed elsewhere form appears in the 1PL and 2PL. These forms are interpreted as present tense forms; there is no difference in tense between the 1SG *finisco* and 1PL *finiamo*. However, it appears as if morphologically, these forms are indeed aligned with the non-present tense forms. So how do we account for this fact?

Focusing specifically on this issue, Embick posits that the 1PL and 2PL forms in the present undergo *impoverishment*, bringing these cells in line with the non-past forms. Impoverishment as an operation has been used in many accounts set within the framework of DM in order to account for cells in a paradigm which seem to be underspecified for some feature (cf. Bobaljik 2002, Harley 2008, Kramer 2016, *a.o.*). Such rules delete morphosyntactic features in certain contexts, expressing the notion that those contexts simply do not express those features in that morphological form.³⁰ For the N-PATTERN, if the feature responsible for the present tense were deleted in the 1PL and 2PL, the use of the elsewhere case in these forms is expected.³¹

²⁹One can also phrase the generalization as occurring in non-past tense ([*-past*]) as framed in Embick (2016).

³⁰A simple case to illustrate this are pronominals in Russian (cf. Bobaljik 2002). Despite making a three way distinction in gender in the 3SG singular, the plural is never distinguished morphologically on the basis of gender. Gender is simply irrelevant when it comes to the morphological expression of the plural. We may thus assume that the gender is impoverished in the plural.

³¹Maiden (2018: 245) claims that ‘...the invocation of ‘Impoverishment’ for the tense value is purely *ad hoc*: it would gain in persuasiveness if the first- and second-person plural forms were seen actually to behave in other respects as if they were not ‘[*-past*]’ (for example, by becoming available as past tense forms), but this is never the case in Italian.’ It is not quite clear what Maiden is referring to here. If he means that the 1PL and 2PL of the present indicative and subjunctive are never used as past tense

An impoverishment rule deletes the [+pres] feature on the T node in the 1PL and 2PL. The impoverishment rule can be stated as follows:³²

$$(6) \quad [+pres] \rightarrow \emptyset / \text{ ___ } [+part, +pl] \quad (\text{Altered from Embick 2016: 292})$$

The rule in (6) accounts for the fact that the 1PL and 2PL pattern with the non-present tense forms by removing the feature that otherwise distinguishes present tense forms from the other forms in the paradigm. In this way, the impoverishment rule effectively says that speakers do not express [+pres] in the 1PL and 2PL.

While it is true that the impoverishment rule in (6) only serves to describe this particular instantiation of the putative N-MORPHOME, the crucial purpose of the analysis is to demonstrate that the morphomic distribution can be described using already established mechanisms without making reference to a paradigmatic template. Thus, the morphomic distribution can be described in terms of two well-established mechanisms: vocabulary insertion and impoverishment.³³

forms, this is expected, as they are still interpreted as present tense in terms of the semantics. The impoverishment only affects the morphological expression. Alternatively, if he means that the *forms* used in the 1PL and 2PL of the present are never used as past tense forms, that is true. However, this is expected, as those forms are accompanied by other tense and mood marking morphology. For example, even if [-past] is impoverished in the present, we do not expect the present tense form *andiàmo* to surface for the imperfect *andavàmo*. This is expected as the impoverishment rule does not suddenly supply a [+imperfect] feature, which is what conditions the presence of additional affixes. The crucial point is that all morphemes which are conditioned by [-past] are affected by the impoverishment. The 1PL and 2PL do resemble the past tense forms, in the sense that *and-* is used and *-sc-* is not.

Maiden (2018: 245) also questions why impoverishment could not condition the deletion of the tense feature in the 1PL and 2PL in the past tense allowing for the spell-out of the present tense in these forms. Such a configuration might well be possible, and if it were attested, an impoverishment rule could account for it. However, speakers do not set up such a rule, as their input simply lacks such a distribution.

³²The original formulation has [-past] instead of [+pres]:

$$(i) \quad [-past] \rightarrow \emptyset / \text{ ___ } [+part, +pl] \quad \text{Embick (2016: 292)}$$

It should also be noted that there may be several functionally equivalent ways of writing the environments expressed by the rules.

³³It is possible to state the analysis in relatively theoretical neutral terms: speakers learn a rule in

It should be noted that the 1PL and 2PL are described in (6) as a unified feature set expressed by the combination of [+participant] and [+plural] features. This unification otherwise goes against other descriptions of the N-PATTERN (e.g. Maiden 2003, 2005, 2018), where the 1PL and 2PL are considered to be disjunct cells that are not targetable by a specific feature set. However, the assumption that the 1PL and 2PL can be unified under a [+participant] feature is not crucial to the analysis. The impoverishment rule in (6) can otherwise be stated as affecting two disjunct environments, one targeting the 1PL and the other the 2PL.³⁴ Thus, the fact that the environment of the N-PATTERN is described here as a unified set is not a crucial part of the analysis.³⁵

In sum, impoverishment as an operation deletes the feature responsible for present tense morphology, [+pres] in the account of Embick (2016), in the 1PL and 2PL, conditioning the appearance of default morphology. In the case of Italian, the default morphology involves the stem *and-* and *fini-* without the suffix *-isc-*.

4.2.5 INTERIM SUMMARY

The last few sections have discussed some of the necessary background for understanding the changes affecting the N-PATTERN. As we have seen, the distribution of *-sc-* in many varieties of Romance has developed from its original distribution on the basis of the N-PATTERN, an irregular pattern based on a stress alternation. Here, the 1PL and

which the present tense is not marked in the 1PL and 2PL. The explanatory power comes from the fact that it is not necessary to posit a morphological template.

³⁴The additional assumption that would have to be made on this alternative is that impoverishment rules can apply to disjunct environments, which is made by Embick (2016) elsewhere. For the analysis here, the impoverishment rule can be written as follows:

- (i) [+pres] → ∅ / ____ [+1, +pl], [+2, +pl]

³⁵In fact, it might be the case that for some of the distributions seen in various other Romance varieties, disjunct impoverishment rules may be preferable. In particular cases with inconsistent patterns in the verb ‘to go’ may be explained with multiple impoverishment rules. For the justification of the presence of multiple disjunct rules, see the discussion on the irregular Latin verbs in Chapter 3.

2PL differ in morphological expression from the other cells of the present tense. While some treatments of the pattern assume that the distribution cannot be described in morphosyntactic terms (e.g. Maiden 2018), I accept Embick’s characterization of the distribution, in which the forms in the singular and 3PL simply express the present tense, while the 1PL and 2PL express the elsewhere form, along with the other tenses.

The next section provides a set of biases which influence the interaction between the N-PATTERN and paradigmatic leveling. The subsequent sections illustrate these biases with discussions of the specific analogical changes.

4.3 PROPOSAL

What is clear from a diachronic point of view is that it is no coincidence that diverse types of morphological expression are aligned with the same distribution, i.e. the N-PATTERN. These distributions are clearly able to exert analogical pressure on one another. Maiden’s templatic approach is admirable, since it captures the generalization that entire paradigms can be restructured in reference to an abstract distribution that can affect completely distinct morphological phenomena, like suppletion, stress, phonologically-induced stem changes, as well as the affixation of suffixes like *-sc-* and *-idi-*. On the other hand, it is not clear whether these templates provide any explanatory power beyond just describing the origin of the distribution and the extension to different paradigms in Romance. In particular, the investigation of how leveling can erode morphomic distributions and produce new ones may provide information on how speakers encode the patterns. Identifying the factors involved in the changes to the N-PATTERN can also improve our understanding of leveling in general. In what follows, I investigate how a morpheme can change into another or can degrade completely when

affected by the analogical leveling of alternants.³⁶

Consider the hypothetical morphomic distribution in (7-a):

(7)

a.

MORPHOME:

b.

CHANGE 1:

c.

CHANGE 2:

Let us posit that two descendants inherited the distribution, but leveled one of the alternants in a subset of the cells differently. Thus, (7-b) and (7-c) show two possible changes affecting the morphomic distribution. We may ask ourselves what factors caused the difference in analogical change. In this section, I posit several biases which can affect the change in distribution.

The hypothesis has two parts. The first is that morphomic distributions are upheld wherever usage, and thus frequency, allow speakers to set up the distributions in their mental grammars. Thus, when changes in usage radically affect the frequency of part of a morphomic distribution, we may see changes in the distribution. To illustrate this, consider the development of the six-cell paradigm in (8).

³⁶Crucially, I am interested in examples of morphomic change that are not influenced by phonological change. For some examples of this, see Maiden (2018: 282-288). It is certainly the case that phonological change could cause morphomic patterns to change in ways completely orthogonal to any notion of morphological coherence; one need only look at the creation of the N-PATTERN to see that phonologically-induced change can create new morphomic patterns. Instead, the focus in this chapter is on change that is entirely morphological and is thus in the realm of analogy, rather than sound change. This focus is not to say that the same factors do not ultimately influence phonologically-induced changes to morphomic distributions.

(8) a. STAGE 1:

b. STAGE 2:

Let us assume that some external construction has replaced the 3PL in terms of usage. This usage would drastically affect the input that new learners would encounter; a learner would have significantly less evidence for the non-shaded form in the 3PL. This change could then provide the impetus to replace the form with default morphology, indicated by the shading of the 3PL in (8-b).

The second part of the hypothesis is that when these changes occur, they are biased in such a way to reflect generalizations that learners are making, specifically in terms of morphosyntactic distributions. In other words, when there is pressure to level part of a paradigm, learners will act in such a way to maximize generalizations that can be expressed in morphosyntactic terms. Consider the following schema:

(9) a. STAGE 1:

b. STAGE 2:

The change depicted in this schema involves the 2SG adopting the morphology of the shaded cells. (9-b) shows a unification of the singular, such that the same alternant is used throughout the singular of the paradigm. This can be seen as a change motivated to create well-formed morphosyntactic feature sets. In other words, morphomic (irregular) distributions exist, but there is pressure to eliminate the morphological irregularity in

favor of more regular distributions.

We can summarize these two parts as follows:

- Usage and frequency determine the maintenance of morphomic distributions.
- Learners are biased to change morphomic distributions to fit with natural classes expressed with morphosyntactic features.

4.3.1 JUSTIFICATION FOR ASSUMPTIONS

While I evidence this hypothesis with discussion of the development of the N-PATTERN in Romance, I provide some initial justifications below.

4.3.1.1 USAGE AND FREQUENCY

Frequency is often assumed to play a role in analogical change. From the early days of modern historical linguistics, frequency has played either an explicit role (Wheeler 1887: 21) or an implicit one (e.g. Paul 1886). Likewise, much research since then has focused on the connection between analogy and frequency (Bybee 1988, 1991, Fertig 1998, Phillips 2006, *a.o.*).³⁷ In this sense, the assumption that the interaction between leveling and morphomic distributions could be related with frequency is uncontroversial.

Special attention can be given to the role of frequency in terms of learning morphological patterns with specific attention to the literature on morphological productivity. Recent research on productivity from synchronic and diachronic perspectives has emphasized the role of frequency (e.g. Baayen 1989, 2009, Sandell 2015, *a.o.*). Particularly notable is the research program of Yang (2016), which sets forth an acquisition model that is based on type frequencies. This model has been applied to diachronic change in-

³⁷cf. Fertig (2013) for more thorough overviews of analogy and its intersection with frequency.

volving analogy, e.g. Yang (2016) for English irregular past tense forms, Kodner (2019, 2020) for Proto-Germanic strong verbs, and Kodner (2022) for Latin past participles.³⁸

Given that much research based around analogy either assumes or directly references type or token frequency, it seems reasonable to posit that frequency may influence the acquisition of morphomic patterns, and through this, the retention and degradation of such patterns. Presumably similar principles can be applied to morphomic distributions as has been to other morphological phenomena, like class distinctions and affixal choice in derivational morphology. It would be odd if morphomic distributions were somehow unique among types of morphological phenomena, in the sense that they would be immune to the influence of frequency and other acquisitional factors.

There is one key part of usage which I see as important for this hypothesis, namely how usage impacts frequency. If there are changes in terms of how parts of the paradigm are used, it is only natural that the frequency in the input for a learner will change as well (when compared to a previous generation). In this way, changes in usage may impact the maintenance of morphomic distributions. As we will see shortly, there is reason to believe that usage changes are behind at least some of the changes in leveling that affect the N-PATTERN.

4.3.1.2 NATURAL CLASSES OF MORPHOSYNTACTIC FEATURES

The assumption that learners attempt to classify the distributions they encounter via natural classes consisting of morphosyntactic features may perhaps be more controversial than the previous assumption.³⁹ However, I hope to demonstrate in this section

³⁸The introduction of default morphology to replace irregular forms can be easily motivated in terms of production if usage impedes the acquisition of irregular morphology in a given cell.

³⁹By ‘morphosyntactic’ features, I refer to the set of properties that express syntactico-semantic concepts. For example, for a verb, we might have features that correspond to different moods and tenses.

that it is worth considering whether there is a diachronic bias towards learning morphological patterns in a way consistent with the morphosyntactic inventory of a given language.

The best argument in favor of this assumption is typological. Particularly illuminating is the typological study of suppletion in [Veselinova \(2006\)](#). Here, one metric according to which suppletion can be classified is the difference between *categorical* and *non-categorical* suppletion. Categorical patterns refer to cases when suppletive forms align themselves along ‘natural’ categories, like tense, aspect, person, or number, while non-categorical distributions cannot be described with natural classes. For our purposes, the non-categorical distributions serve as a proxy for morphomic distributions, which likewise cannot be described with natural classes. The relevant insight that we can draw from [Veselinova \(2006\)](#) is that while non-categorical patterns of suppletion for verbs exist, these examples are generally more restricted than categorical patterns and are often found with auxiliaries and other highly irregular verbs rather than regular lexical verbs ([Veselinova 2006](#): 83-86). This tendency can be explained by the fact that these idiosyncratic environments are not only much more frequent, but also, in the terms of [Bybee \(1985\)](#), autonomous in the sense that the forms cannot be derived from other forms and must be memorized directly. These factors together allow for the maintenance of irregular suppletive patterns.

I argue that the diachronic biases that force suppletion to largely cohere in terms of natural classes can similarly be applied to other morphological distributions. However, the comparison between suppletive and morphomic distributions does not only serve as a parallel as the two are not completely distinct from one another. As suppletion plays a key role in most morphomic distributions, the principles that produce the categorical and non-categorical patterns for suppletion also intersect with morphomic distributions.

The assumption that learners prefer morphological phenomena to mark natural categories also explains the opposition between the morphological marking of the present and of the non-present tenses seen in Romance. As discussed in Section 4.2.4, the N-PATTERN largely coincides with the marking of the present tense. The leveling of alternants marking the 1PL and 2PL in favor of the alternant which marks the present tense reinforces that these forms are present tense forms.

If we take this point together with the distributions summarized in Section 4.1.1, a diachronic trend becomes clear. The changes with regard to intra-paradigmatic leveling serve to largely reinforce the difference between the present and non-present tenses. In both the pathways, the general trend is to eliminate the alternations within the present, and crucially not to eliminate the alternation between the present and non-present tenses. By assuming that speakers have a bias to associate morphological marking with a natural class, we can motivate the general trend in leveling.

Admittedly, the leveling that variably affected the N-PATTERN does not prove that speakers have a bias in and of itself. However, it is worth considering whether leveling in general is influenced by such considerations as it provides a motivation for the analogical change that goes beyond description.

4.3.2 FRAMING PRELIMINARIES

Before discussing the specifics of each stage of development, I briefly summarize what the pathways are and discuss a categorization of the different stages that make up the pathways. Recall the two pathways presented in Section 4.1.1 regarding the distribution of *-sc-*, restated below:

(10) PATHWAY 1

a. Stage 1: N-PATTERN in Lombard (Mantuan)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

b. Stage 2: 2PL CONSERVING in Lombard (Cremonese, Bergamasque, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

c. Stage 3: 2PL IND CONSERVING in Lombard (Milanese, Bresciano, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

(11) PATHWAY 2

a. Stage 1: N-PATTERN in Catalan

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

b. Stage 2: 1/2PL IND CONSERVING in Ligurian, Old Occitan, etc.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

c. Stage 3: 1PL IND CONSERVING in Piedmontese

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

Each pathway demonstrates a different set of stages in which a cell containing the $\emptyset$ -alternant has been eliminated in favor of the alternant containing *-sc-*. I suggest that the patterns of elimination described above can be broken down into two basic changes, i) introduction of *-sc-* to the entire subjunctive, and ii) introduction of *-sc-* to a specific person + number combination (e.g. extension of *-sc-* to the 1PL), shown in (12) by blue and red respectively.

- (12) a. PATHWAY 1: N-PATTERN
- *EXTENSION TO 1PL*
- *LEVELING IN THE SUBJUNCTIVE*
- b. PATHWAY 2: N-PATTERN
- *LEVELING IN THE SUBJUNCTIVE*
- *LEVELING OF 2PL*

The justification for this separation is twofold. Firstly, by assuming that the motivation behind the leveling in the subjunctive in both pathways can be unified, we have the potential of finding a more powerful explanation for why the forms in the subjunctive are more prone to level before forms in the indicative. While I concede that the leveling in the subjunctive in each of the pathways may have occurred independently for completely idiosyncratic reasons, an explanation in which the leveling in the subjunctive can be unified is preferable.

Secondly, it is also worth considering the hypothesis that intra-paradigmatic leveling like that in the subjunctive is ultimately subject to different considerations than leveling across multiple tenses, which accompanied the leveling of person + number (that is, the elimination of the -Ø- alternant). The development in the first pathway illustrates this difference. The leveling of the subjunctive is internal to its local paradigm; it did not cause leveling in the indicative. The extension of -sc- to the 1PL, however, occurred in both the indicative and subjunctive, and therefore was not restricted to its local paradigm.

TERMINOLOGICAL NOTE: Before continuing, it is necessary to make a small clarification regarding the terminology used to describe the different distributions. In (10) and (11) above, I have described the distributions derived from the N-PATTERN with various

labels. Moving forward, I will keep calling the general distribution seen in (10-a) and (11-a) as the N-PATTERN. The various derivatives of this pattern are labeled in terms of the aberrant *non-leveled* cells. Thus, the N-PATTERN distribution in (11-b) where the subjunctive has leveled is called 1/2PL IND CONSERVING, meaning that the 1ST PL and 2ND PL forms of the present *indicative* conserve the old pattern. The same naming schema has been applied to the other patterns. PATHWAY 1 and PATHWAY 2 can be stated thusly:

- (13) a. PATHWAY 1: N-PATTERN $\rightarrow$ 2PL CONSERVING $\rightarrow$ 2PL IND CONSERVING
 b. PATHWAY 2: N-PATTERN $\rightarrow$ 1/2PL IND CONSERVING $\rightarrow$ 1PL IND CONSERVING

With this in mind, we can move on to discussing the individual stages of development.

4.4 EXTENSION OF -SC- BASED ON PERSON + NUMBER

I first consider the extension of -sc- in terms of person and number before moving on to the leveling specific to the subjunctive (Section 4.5).

4.4.1 INTRODUCTION OF -SC- TO THE 1PL

This section discusses several Lombard dialects which have introduced -sc- to the 1PL. Section 4.4.1.1 gives an overview of the data. Section 4.4.1.2 and Section 4.4.1.3 provide evidence for the fact that usage and frequency have caused the extension of -sc- to the 1PL.

4.4.1.1 DATA FROM LOMBARDY (2PL CONSERVING)

The Gallo-Italic dialects of Lombardy exhibit distributions based upon the above pattern. In the eastern dialects of Lombard, we largely find the original N-PATTERN distribution throughout the present indicative and present subjunctive, parallel to Italian. For example, consider the Mantuan and Gavardo (both Lombard dialects) paradigms in Table 4.13 and Table 4.14 in which the stressed reflex of *-sc-* is found in the N-PATTERN:⁴⁰

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>polìsi</i>	<i>polìsi</i>	<i>polìs</i>	<i>polém(a)</i>	<i>polì</i>	<i>polìs</i>
PRES SUBJ	<i>polìsa</i>	<i>polìsi</i>	<i>polìsa</i>	<i>poléma</i>	<i>polìghi</i>	<i>polìsa</i>

Table 4.13: PRES IND/SUBJ of *polir* ‘to clean’ in Mantuan (Agarotti 2005: 279)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>fenesè</i>	<i>feneset</i>	<i>fenes</i>	<i>fenóm</i>	<i>finí</i>	<i>fenes</i>
PRES SUBJ	<i>feneses</i>	<i>feneset</i>	<i>feneses</i>	<i>fenómes</i>	<i>finíef</i>	<i>feneses</i>

Table 4.14: PRES IND/SUBJ of *finir* ‘to finish’ in Gavardo (Alberti 1993: 97-105)

The general stress pattern in the language also conforms to the N-PATTERN, as in Table 4.15 and Table 4.16.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>córi</i>	<i>córi</i>	<i>cór</i>	<i>corém(a)</i>	<i>corì</i>	<i>cór</i>
PRES SUBJ	<i>córa</i>	<i>córi</i>	<i>córa</i>	<i>coréma</i>	<i>corìghi</i>	<i>córa</i>

Table 4.15: PRES IND/SUBJ of *córar* ‘to run’ in Mantuan (Agarotti 2005: 277-278)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>ríde</i>	<i>rídet</i>	<i>rít</i>	<i>ridóm</i>	<i>ridí</i>	<i>rít</i>
PRES SUBJ	<i>rídes</i>	<i>rídet</i>	<i>rídes</i>	<i>ridómes</i>	<i>ridíef</i>	<i>rídes</i>

Table 4.16: PRES IND/SUBJ of *ríder* ‘to laugh’ in Gavardo (Alberti 1993: 97-105)

⁴⁰The vowel change in the 2PL in Gavardo is due to umlaut caused by the vowel *í* (Alberti 1993: 97).

These paradigms preserve the regular distribution of the N-PATTERN seen in Italian, Catalan, and Romanian. Mantuan and Gavardo therefore represent the original distribution in Lombard. As we will see, however, many Lombard dialects have shifted to different distributions

In some dialects, the 1PL of both the present indicative and present subjunctive has been ‘regularized’ to more resemble the forms in the singular and 3PL.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisi</i>	<i>finiset</i>	<i>finis</i>	<i>finisum</i>	<i>finii</i>	<i>finis</i>
PRES SUBJ	<i>finisa</i>	<i>finisa</i>	<i>finisa</i>	<i>finisum</i>	<i>finii</i>	<i>finisa</i>

Table 4.17: PRES IND/SUBJ of *finii* ‘to finish’ in Cremonese (Taglietti 2005: 252)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>capèsse</i>	<i>capèsset</i>	<i>capèss</i>	<i>capèss</i>	<i>capì</i>	<i>capèss</i>
PRES SUBJ	<i>capèsse</i>	<i>capèsset</i>	<i>capèsse</i>	<i>capèsse</i>	<i>capì</i>	<i>capèsse</i>

Table 4.18: PRES IND/SUBJ of *capì* ‘to understand’ in Bergamasque (Zanetti 2005: 172)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisi</i>	<i>finisät</i>	<i>finisä</i>	<i>finisäm</i>	<i>finì</i>	<i>finisän</i>
PRES SUBJ	<i>finisä</i>	<i>finisät</i>	<i>finisä</i>	<i>finisäm</i>	<i>finì</i>	<i>finisän</i>

Table 4.19: PRES IND/SUBJ of *finì* ‘to finish’ in Paves (Comoletti 2005: 71)

The difference between this distribution and the regular N-PATTERN is that only the 2PL of both the present indicative and present subjunctive lacks the suffix *-is*, rather than both the 1PL and 2PL. This new pattern is also reflected by the rest of the verbal system in these languages. For example, the stress pattern of the other conjugations fit the same distribution in Cremonese (Table 4.20) and Bergamasque (Table 4.21):⁴¹

⁴¹The double vowel sequences in Cremonese and in other dialects refer to long vowels. Thus 1SG *pàarli* contains a long stressed [a:], while 2PL contains a long stressed [e:]. The stress patterns are often correlated with changes in vowel quality, although the exact alternation caused by the stress pattern is dependent on the dialect in question.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>pàarli</i>	<i>pàarlet</i>	<i>pàarla</i>	<i>pàarlum</i>	<i>parlée</i>	<i>pàarla</i>
PRES SUBJ	<i>pàarla</i>	<i>pàarla</i>	<i>pàarla</i>	<i>pàarlum</i>	<i>parlée</i>	<i>pàarla</i>

Table 4.20: PRES IND/SUBJ of *parlàa* ‘to speak’ in Cremonese (Taglietti 2005: 252)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>crède</i>	<i>crèdet</i>	<i>crèd</i>	<i>crèd</i>	<i>credi</i>	<i>crèd</i>
PRES SUBJ	<i>crède</i>	<i>crèdet</i>	<i>crède</i>	<i>crède</i>	<i>credighev</i>	<i>crède</i>

Table 4.21: PRES IND/SUBJ of *crèd* ‘to believe’ in Bergamasque (Zanetti 2005: 170)

It is clear that this restricted distribution originally stems from the N-PATTERN. The conserved cells in the present are a subset of those found in the N-PATTERN, suggesting a straightforward restriction of aberrant cells. This is also supported by the fact that some dialects, like Cremonese, preserve the N-PATTERN in some verbs, e.g. the present tense for *andàa* ‘to go’ (cf. Section 4.4.1.3). Given that it is unlikely that the N-PATTERN in these irregular verbs stems from the more restricted pattern, we can assume that the latter is an innovation in these dialects based on an original N-PATTERN distribution.

4.4.1.2 DIACHRONIC EXPLANATION:

The traditional explanation for the notable intrusion into the 1PL is that the 1PL has been largely replaced with an impersonal construction based on the 3SG (Meyer-Lübke 1894: 168) throughout the different conjugations, specifically in the dialects which have done this leveling. Thus, in Bergamo and some of the eastern dialects of Lombard, we find the following construction: Berg. *am kanta* ‘one sings’ from Latin *homo cantat* ‘one sings’, having replaced the reflex of Latin *cantamus* ‘we sing’ in terms of usage. This usage is comparable to the replacement of the 1PL *nous* ‘we’ with *on* ‘one’ in spoken French (Flechia, 1876, Grafström, 1969, *a.o.*).

Meyer-Lübke (1894: 168) proposes that this usage ultimately produced a *Mischform* in the 1PL, which was present throughout all of Lombardy. Thus, the Milanese form

kántem ‘we sing’, for the 1PL, with stress on the root, rather than the desinence, was the outcome of the mixing of the older forms derived Latin *cantámus* where stress was on the theme vowel and the 3SG *am kánta* (used for the 1PL) where stress is on the root. The stress placement in this newly created form occurred in all regular conjugations, producing a stress pattern in which every cell in the present indicative and subjunctive apart from the 2PL shows stress on the root. As for the forms with *-is(s)-* < Lat. *-sc-*, the suffix, which hosted stress on the preceding vowel, was extended to the first person as well, conforming with the newly emergent stress pattern.

This explanation is not without its critics. Huber (1958), in his study of the dialect of Livigno, finds issue with the claim that the root stress is due to a ‘mixing’, as there is no evidence that learners had simultaneous input of the 1PL (without root stress) and the impersonal 3SG construction which would result in a *Mischform*. Additionally, Huber challenges the assumption that the *am kanta*-construction was ubiquitous throughout the entire Lombard region, as many of the dialects, including Livigno, lack this construction, and its prehistoric presence in the region is doubtful. Indeed, it is peculiar that at least some eastern dialects do not replace the 1PL, while the western dialects introduced *-sc-* to the 1PL across the board.

Instead of referencing the impersonal construction, Huber (1958) proposes an alternative account in which stress was introduced to the 1PL in the first, second, and fourth conjugations in accordance with the stress pattern in the third conjugation, where the stress normally falls on the root. The resulting forms underwent regular phonological developments producing the attested forms, e.g. Liv. 1ST CONJ. *pórtom* ‘we carry’, 4TH CONJ. *dórmom* ‘we sleep’. In other words, the stress in the preforms **portámus*, **dormímus* was moved to the root syllable, in accordance with the third conjugation, cf. **véndimus* ‘we sell’ > *véndom*. These newly stressed forms underwent vowel reduc-

tion, before the final syncope produced the attested forms: **pórtamus* > **pórtǔmus* > *pórtom*, **dórmimus* > **dórmǔmus* > *dórmom*.⁴²

While the issues that Huber (1958) brings up against the account proposed by Meyer-Lübke (1894) hold some weight, in my view, his alternative account presents even more significant problems. The issues that this account poses are twofold. Firstly, it is difficult to motivate an analogical shift of the stress on the 1PL from the third conjugation to the other conjugations when the other conjugations would have far outnumbered the verbs in the third conjugation. Indeed, this directly contradicts the general trend in Romance to eliminate the root stress from the third conjugation (cf. Meul 2016, 38-40). Secondly, it is not clear why stress would be shifted in the 1PL, but not the 2PL. For each of the conjugations in Livigno (as well as the other Lombardy dialects), *including* the third conjugation, the 2PL is not stressed on the root, e.g. *vendé* ‘you (PL) sell’.⁴³ This is expected, given that the vast majority of Romance languages have shifted stress away from the root in the 1PL and 2PL of the third conjugation. However, the proposed development of stress in Livigno by Huber (1958) supposes that the third conjugation kept the stress on the root in the 1PL, and even exerted influence on the other conjugations. If this was the case, why did this retention not occur for the 2PL? In other words, for Huber (1958), the third conjugation serves as an attractor for the other conjugations in the 1PL, but is influenced by the other conjugations for the 2PL. Such a development is simply unmotivated both internally and typologically.

Taken together, it is unlikely that the stress of the 1PL of the first, second, and fourth conjugations was aligned to be more similar to the third conjugation, but the exact opposite change affected the 2PL.

⁴²Huber (1958) posits that the short unstressed vowels became a rounded vowel before a labial.

⁴³The dialects of Lombardy that are 2PL IND CONSERVING (e.g. Milanese) do show a present subjunctive which is stressed on the root; however, as will be discussed later, this is a later development and is not based on an archaic retention of the 2PL form from the third conjugation.

To reconcile the explanation proposed in Meyer-Lübke (1894) with the concerns raised by Huber (1958), it may be prudent to reframe the change. Instead of the change being due to a ‘mixing’ of the 3SG and 1PL, we can assume that the 1PL forms were produced productively, based on the regular pattern in which most present tense forms were created. Given that a stressed stem (e.g. *kánt-*) would have been associated with most of the present tense (i.e. all of the indicative and subjunctive apart from the 1PL and 2PL), a productively created form would use this stem. This productive use of the stem would be motivated by the assumption that if the usage in these dialects was such that learners were deprived input for the 1PL, it is likely that they would produce a form in accordance with the general pattern of the present tense, i.e. a form with a stressed root. As the use of the *am kanta* construction would have limited the input of the 1PL, the old 1PL was replaced with a regularized form of the stem that was stressed in accordance with the rest of the paradigm. Likewise for verbs without the suffix *-íss-*, learners might have posited a form of the stem that was in accordance with most of the present tense, thus introducing the *-sc-* suffix to the 1PL. In this sense, the change in the verbs with *-sc-* exactly paralleled the changes in the other conjugations, without reference to stress.⁴⁴

The fact that this change did not affect the 2PL is expected if the stress shift was formed productively, as the input for the 2PL would not have been affected by the usage of the 3SG impersonal construction in lieu of the 1PL. Thus, learners would simply

⁴⁴There is also the possibility raised by Huber (1958) that ‘mixing’ might refer to an interpretation in which the pronoun *am* < Lat. *homo* was added as a suffix to the 3SG *kánta*, to produce a new 1PL form. Huber (1958) dismisses this possibility as unlikely, given that the pronoun never follows the verb. I leave open whether this hypothesis is a viable alternative to the proposal I posit here, although it should be noted that there is little evidence in support of such a change. With regards to the stem *kánt-*, it is not clear to me whether there is a meaningful difference between deriving its use in the 1PL directly from the 3SG as opposed to deriving it from a generalized present stem. This hypothesis does provide a source for the ending, although the nature of the ending is orthogonal to the leveling with which we are concerned with here.

maintain the pattern in the input, namely an unstressed root syllable.

As for the issue raised by Huber (1958) that leveling in the 1PL occurs in dialects that never had the impersonal construction, I suggest that the change began in those dialects that had the replacement of the 1PL with the 3SG, before diffusing through the well-established dialect continuum (cf. Sanga 1997: 253).⁴⁵ In this sense, Lombard dialects like Livigno did not productively produce the forms in the 1PL; rather they replaced the non-root stressed forms with stressed forms as a result of contact with other dialects. It was therefore not necessary for each dialect to remodel the 1PL independently.⁴⁶

In sum, I believe the general assumption maintained by Meyer-Lübke (1894), Rohlf (1949), Meul (2016), *a.o.* that the forms in the 1PL were remodeled in accordance with the rest of the paradigm due to the usage of the impersonal construction to be justified. The key observation is that the introduction of *-sc-* to the 1PL likely only occurred because a change in usage severely restricted learners' input, allowing for the production of a 'more regular' form. There is no evidence for a paradigm internal motivation for the leveling in terms of an 'output-to-output' constraint (cf. Michael 1996, Burzio 2005, McCarthy 2005, *a.o.*). Nor does the leveling in this paradigm immediately suggest that there is a preference to level the 1PL instead of the 2PL in alternations with distributions like the N-PATTERN. Rather, without the external motivation (the usage of the impersonal construction for the 1PL), these dialects likely would have retained the N-PATTERN distribution.

⁴⁵While the individual dialects that make up the dialect continuum are currently distinct (Sanga 1997, Benincà 2000), we can assume that at the time of the creation of the 1PL form, there was an opportunity for diffusion: "there was a time when a basically Milanese type extended from Novara up to Trento and across to Verona; and it is this common base that constitutes the unity of the Lombard dialect group, whereas differences are due to later more superficial developments" (Sanga 1997: 253).

⁴⁶Future research may investigate whether the dialect variation in terms of the 1PL demonstrates such a contact situation.

4.4.1.3 SUPPLETIVE INCONSISTENCIES

In many of the Lombard dialects that have introduced *-sc-* to the 1PL, there exist patterns that do not cohere to the expected pattern of inflection for the suppletive paradigm of the verb ‘to go’. For example, in Cremonese (a dialect which regularly shows a 2PL CONSERVING distribution), the verb still maintains the N-PATTERN in its distribution, with *’nd* forms in the 1PL and 2PL and *va* forms elsewhere:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>vòo</i>	<i>vèet</i>	<i>và</i>	<i>’ndóm</i>	<i>’ndèe</i>	<i>và</i>
PRES SUBJ	<i>vàga</i>	<i>vàga</i>	<i>vàga</i>	<i>vàagum/endèm</i>	<i>endèe</i>	<i>vàga</i>

Table 4.22: PRES IND/SUBJ of *andàa* ‘to go’ in Cremonese (Taglietti 2005: 254)

As can be seen in Table 4.22, the singular and 3PL show the *v-* < *vad-* stem throughout the present indicative and present subjunctive, while the *end-* < *and-* appears elsewhere.

Another example can be seen in Milanese, which usually shows a 2PL IND CONSERVING distribution:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>voo</i>	<i>vee(t)</i>	<i>va</i>	<i>vèmm/andèmm</i>	<i>vii/andii</i>	<i>vann</i>
PRES SUBJ	<i>vaga</i>	<i>vaghet</i>	<i>vaga</i>	<i>vagom</i>	<i>vagov/andii</i>	<i>vaghen</i>

Table 4.23: PRES IND/SUBJ of *andà* ‘to go’ in Milanese (Nicoli 1983: 310)

Where we might expect a *v-* stem in the 1PL of the indicative (as well as a fully leveled subjunctive), instead we find a pattern inconsistent with the rest of the verbal system.⁴⁷

These paradigms, despite being inconsistent when considering the verbal system as a whole, are perfectly amenable to an explanation that assumes that morphomic

⁴⁷The nature of the variation seen in these paradigms is not clear, as the grammars are not particularly forthcoming. Regardless of what the variation is, the important fact is that there seems to be a discrepancy between the verb ‘to go’ and other verbs.

The distributions in both Cremonese and Milanese seem to suggest that the 1PL of the subjunctive is more likely to level than the 1PL of the indicative. We will return to this point in Section 4.5.

distributions are affected by the usage and frequency. While there is little information on the usage and frequency of these dialects, the usage of the impersonal phrase for the 1PL provides key insight. Coveney (2000) examines the distribution of *nous* ‘we’ and *on* ‘one’ in informal spoken French in several speakers. A relevant finding is that although every verb engages in the replacement of the 1PL with the 3SG, the use of *nous* primarily occurs with a small set of highly frequent, irregular verbs, including *avoir* ‘to have’, *être* ‘to be’, and *aller* ‘to go’.⁴⁸ This fact indicates that the older usage is correlated with verbs of this type. For the Lombard dialects, we have to assume that a similar development occurred, where verbs like ‘to go’ simply preserve the old distributions more frequently than regular verbs.⁴⁹

4.4.1.4 INTERIM SUMMARY

In sum, the introduction of *-sc-* to the 1PL, and thus the elimination of the $-\emptyset$ - alternant in this cell, is motivated by a usage-based change in frequency. An impersonal construction based on the 3SG limited the input for this cell, allowing learners to productively create a new 1PL. This analysis is corroborated by the fact that the suppletive alternations involving the verb ‘to go’ do not show a replacement of the 1PL, as motivated by the special usage of the 1PL of this verb. Learners would have had input for this

⁴⁸cf. Coveney (2000: 475): “The two most frequent verbs [*avoir* and *être*] in French together account for a large proportion of the 49 occurrences of *nous* + 4p [i.e. 1PL] verb. Thirteen other verbs figure in the total, including the irregular verbs *aller* (‘to go’), *connaître* (‘to know’), *faire* (‘to do’), *recevoir* (‘to receive’) and *repandre* (‘to take back’), which suggests that use of the 4p verb now tends to be restricted in informal French to a limited number of common irregular verbs.”

⁴⁹Interestingly, there are also inconsistencies in Cremonese between forms with stress on the root and forms with stress on the suffix *-um* for the 1PL. Rossini (1975: 76-77) notes that, while the more recent form (with stress on the root) is more frequent, the older form is still found especially as exhortative subjunctives (‘let us X’) and whenever there are enclitic particles. This is also the case with the form *finúm*, which appears instead of *finisum*. However, there are also some *-is-* verbs that do not show any variation: *stremisum* ‘we are frightened’ but not **stremúm* (Rossini 1975: 77). This distribution finds another parallel with the French case, where imperatives in the 1PL (i.e. exhortative subjunctives) often retain the use of the regular 1PL form: *préparons-nous* ‘let us prepare ourselves’ (Coveney 2000: 461).

specific form, allowing for the maintenance of the N-PATTERN in this environment.

We can formalize this change with a modified rule based on (6). Here, the environment has been restricted to just the 2PL (from the 1PL and 2PL):

$$(14) \quad [+pres] \rightarrow \emptyset / \text{---} [+2, +pl]$$

2PL CONSERVING in Lombard (Cremonese, Bergamasque, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

4.4.2 INTRODUCTION OF -SC- TO THE 2PL

This section discusses how the suffix *-sc-* was introduced to the 2PL in Piedmontese. Section 4.4.2.1 begins by outlining the distribution in this dialect. Section 4.4.2.2 provides a diachronic explanation for the extension based on frequency.

4.4.2.1 DATA FROM PIEDMONTESE (1PL IND CONSERVING)

Piedmontese demonstrates another change involving the introduction of *-sc-* based on person and number. Here, the suffix has been extended to the 2PL. The consequence of this change is that the N-PATTERN has been restricted to the 1PL of the present indicative.⁵⁰

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisso</i>	<i>finisse</i>	<i>finis</i>	<i>finioma</i>	<i>finisse</i>	<i>finisso</i>
PRES SUBJ	<i>finissa</i>	<i>finisse</i>	<i>finissa</i>	<i>finisso</i>	<i>finisse</i>	<i>finisso</i>

Table 4.24: PRES IND/SUBJ of *finí* ‘to finish’ in Piedmontese (Grosso 2002: 96-97)

As seen in Table 4.24, *-is(s)-*, the reflex of Latin *-sc-*, is found throughout the present

⁵⁰This dialect has also leveled the subjunctive in favor of the alternant with *-sc-*. For this, see Section 4.5.

tense apart from in the 1PL present indicative. This same pattern is replicated by the general stress pattern:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>cánto</i>	<i>cánte</i>	<i>cánta</i>	<i>cantóma</i>	<i>cánte</i>	<i>cánto</i>
PRES SUBJ	<i>cánta</i>	<i>cánte</i>	<i>cánta</i>	<i>cánto</i>	<i>cánte</i>	<i>cánto</i>

Table 4.25: PRES IND/SUBJ of *canté* ‘to sing’ in Piedmontese (Grosso 2002: 90-91)

From a diachronic perspective, the Piedmontese distribution represents a meaningful adaptation of the pattern seen in Section 4.5, in that the subjunctive has been leveled in favor of the *-sc-* alternant. The only difference is that the 2PL of the PRES IND has also been affected. This distribution further supports the trend for Romance languages to reduce the anomalous cells in the N-PATTERN in order to maximize the marking of the present tense.

4.4.2.2 DIACHRONIC EXPLANATION:

The general chronological development in Piedmontese is clear. The more restrictive pattern, namely the 1PL IND CONSERVING pattern, was derived from a prior 1/2PL IND CONSERVING pattern, as confirmed by the paradigm for ‘to go’, which preserves the 2PL indicative *and-* stem (*andève*) while maintaining a leveled subjunctive pattern (see Table 4.27 below). Thus, the restriction of the 2PL should be seen as a later innovation than the leveling of the subjunctive. However, exactly how the suffix *-sc-* was extended to the 2PL of both the PRES IND and PRES SUBJ in Piedmontese is not immediately obvious and, as we will see, poses an interesting case study of leveling. This section addresses the introduction of *-sc-* to the 2PL, while the leveling of the subjunctive is addressed in Section 4.5.

It is important to note that in most verbal paradigms in Piedmontese, the 2SG and 2PL are identical in form. Recall the paradigm in Table 4.24, restated below in

Table 4.26:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisso</i>	<i>finisse</i>	<i>finis</i>	<i>finioma</i>	<i>finisse</i>	<i>finisso</i>
PRES SUBJ	<i>finissa</i>	<i>finisse</i>	<i>finissa</i>	<i>finisso</i>	<i>finisse</i>	<i>finisso</i>

Table 4.26: PRES IND/SUBJ of *finí* ‘to finish’ in Piedmontese (Grosso 2002: 96-97)

The form *finisse* represents the outcome for the second person in both the plural and singular in the subjunctive (as well as the indicative).

For most verbs, the homophony of the 2SG and plural is found throughout different tenses or moods, although there are exceptions, suggesting that there is a general meta-syncretism between the 2SG and 2PL.⁵¹ For example, the forms of the future are accompanied by different endings in the singular and plural, e.g. 2SG *finiras* ‘you will finish’ versus 2PL *finireve* ‘you will finish’ (Grosso 2002: 97). Irregular verbs show these endings in both the present indicative and future, showing a lack of homophony between the 2SG and 2PL in these tenses. Particularly enlightening is the paradigm for the verb *andé* ‘to go’, which not only shows a difference in desinence but also in stem between the 2SG and 2PL:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>vad</i>	<i>vas</i>	<i>va</i>	<i>andoma</i>	<i>andeve</i>	<i>van</i>
PRES SUBJ	<i>vada</i>	<i>vade</i>	<i>vada</i>	<i>vado</i>	<i>vade</i>	<i>vado</i>

Table 4.27: PRES IND/SUBJ of *andé* ‘to go’ in Piedmontese (Griva 1980: 84)

That said, the generalization still stands that for the vast majority of paradigms, the 2SG and 2PL are identical in form.

In terms of diachrony, the general explanation for the concordance between the

⁵¹A meta-syncretism refers to syncretism that occurs across multiple words and morphological contexts. In Piedmontese, the 2SG and 2PL are syncretic in the vast majority of cases, suggesting that there is a meaningful generalization at play. See Williams (1994), Bobaljik (2002), Harley (2008), Kramer (2016) for more details on meta-syncretisms.

The paradigm in Table 4.27 shows that the syncretism between the 2SG and 2PL is not present throughout the entire language, suggesting that the forms 2SG *finisse* and 2PL *finisse* despite being identical on the surface do have a different underlying structure.

2SG and 2PL forms is that the desinence *-e* seen in both is the phonologically expected outcome of the respective analogically remodeled forms. Beginning with the 2SG, the *-e*, found in all conjugations, is the regular outcome of earlier *-es*. In many Italian dialects, including Standard Italian, the 2SG ending *-as* of the first conjugation was replaced by the ending of other conjugations, although the exact nature of the replacement is dependent on the dialect. Thus, the 2SG *-e* seen in Piedmontese is likely the reflex of the second conjugation ending Lat. *-ēs* (Rohlf 1949: 289). Other northern dialects have instead taken the *-i* from the fourth declension (Rohlf 1949: 289), e.g. Mantuan 2SG *polìs-i* < *-īs*.⁵²

The 2PL *-e*, on the other hand, can be derived from the first conjugation Lat. *-ātis*, by means of *-ade*, *-ai* > *-e*. The northern dialects that otherwise have *-i*, *-ii* [i:], etc. derive their endings from the second and fourth conjugations (Rohlf 1949: 296). Together, the analogical changes affecting these endings can be seen to be the result of a general conflation between the different conjugations.⁵³

The endings seen in the irregular verbs and the future are straightforward. The ending *-s* in the 2SG was maintained word-finally in these monosyllables, e.g. Piedmontese *stas* ‘you (2SG) stand’ < Lat. *stās*. The ending *-eve* is likely simply the addition of pronominal material from the 2PL pronoun *voi* to the regular ending *-e*. Similar developments have also occurred throughout Lombardy: Milanese *capíssov*, Bergamo *credìghev*, etc.

Though we are still unable to derive the distribution of *-sc-* in Piedmontese through these developments alone, as we would expect **finé* as the reflex of the 2PL instead of *finísse*, one might imagine the following scenario. The phonological identity between

⁵²Dialects in this region often add pronominal material to reinforce the endings, in particular the 2SG and plural, e.g. Milanese 2SG *-et* < *-e* + *te*.

⁵³See Appendix Appendix B.1.2.

endings of the 2SG and 2PL indicates that learners would have encountered a pattern in which the second person forms throughout the vast majority of the verbal system were identical. We merely have to assume that learners acquired an abstract rule in which the 2PL was identical in form with the 2SG. Likely beginning with highly productive first conjugation verbs (e.g. 2SG/2PL *cánte*), the extension of this rule to the *-iss* type would have produced a 2PL *finísse*, replacing an earlier **finé*.⁵⁴

Why was the N-PATTERN not preserved in this instance? I argue that frequency—and specifically the low frequency of the 2PL compared to other persons and numbers—plays a role in the relative lack of maintenance of the morphomic distribution.⁵⁵ In addition to being generally considered the most marked member of a six-celled paradigm (cf. [Harley & Ritter 2002a,b](#), *a.o.*), the 2PL has been shown to be the least frequent person and number combination. It has, for example, been shown for both written ([Juilland & Chang-Rodríguez 1964](#)) and spoken Spanish ([Rodríguez Bou et al. 1952](#)) that the 2PL is the least common of each person + number combination in multiple tenses.⁵⁶ That the 2PL is learned last has subsequently been confirmed by bilingual studies in Basque and Spanish ([Ezeizabarrena 1996](#), [Austin 2013](#)) as well as a monolingual study in Spanish ([Aguirre, 2003](#)).⁵⁷ Even though I do not have any specific information related to the frequency distributions of Piedmontese, it seems reasonable to assume that the 2PL also has a low frequency relative to the other person and number combinations.⁵⁸

There are several characteristics of the morphology of Piedmontese itself that support the notion that the low frequency of the 2PL influenced the extension of *-sc-* (and

⁵⁴There remains the question of why stress was moved from the ending to the preceding syllable (i.e. **canté > cánte*). The exact conditions by which this change has occurred are difficult to ascertain, and I leave it open to further research to expand on exactly why this shift occurred.

⁵⁵Note that the 2PL IMP has a high frequency, as we will see below.

⁵⁶See also [Bybee \(1985, 2006a\)](#).

⁵⁷See also [Pizzuto & Caselli \(1992\)](#) for a similar result for Italian.

⁵⁸However, it should be noted that the methodology of child acquisition studies are often biased against the production of 2PL forms ([Hyams \(1986\)](#) in [Pizzuto & Caselli 1992](#): 504).

the other stems) to the 2PL. The first is that the irregular and likely highly frequent verb *andé* ‘to go’, as seen in Table 4.27, preserves the older pattern, in which the 2PL shows the *and-* stem (*andève* ‘you (PL) go’), rather than the expected *va(d)-* stem. It seems likely that the high frequency of this verb would provide enough input of its 2PL form for learners to block any such analogical change. Such a development would also mirror the development we see in terms of the 1PL described in Section 4.4.1.3, where the 1PL of several highly frequent irregular verbs preserves an older distribution. In this sense, the fact that the verb ‘to go’ shows similar morphomic inconsistencies across different languages can be unified under one explanation; the high frequency of the verb allows for the preservation of older morphomic distributions.

The second relevant aspect of Piedmontese morphology is that the 2PL imperative does *not* show extension of more regular present tense morphology, including the extension of *-sc-*. Consider the full present tense of *finí* which includes the imperative:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisso</i>	<i>finisse</i>	<i>finís</i>	<i>finioma</i>	<i>finisse</i>	<i>finisso</i>
PRES SUBJ	<i>finissa</i>	<i>finisse</i>	<i>finissa</i>	<i>finisso</i>	<i>finisse</i>	<i>finisso</i>
PRES IMP	–	<i>finiss</i>	<i>finissa</i>	<i>finioma</i>	<i>finì</i>	<i>finisso</i>

Table 4.28: PRES IND/SUBJ of *finí* ‘to finish’ in Piedmontese (Grosso 2002: 96-99)

The lack of *-sc-* in the imperative is a remnant of the N-PATTERN distribution, cf. Italian 2PL PRES IMP *finite* ‘finish!’. The preservation of the older pattern in the imperative follows directly from a frequency account, as the 2PL imperative does not have the same relative low frequency in comparison to other person and number combinations as the indicative (see for example Pizzuto & Caselli (1992) on Italian). In this sense, the 2PL imperative was protected from analogical leveling just as the present indicative of *andé* ‘to go’ was.

In summary, the lack of extension of ‘present tense’ morphology to highly frequent morphological environments suggests that frequency influences the maintenance of mor-

phomic patterns. Essentially, if there is simply not enough evidence for certain parts of the morphomic distribution, we may expect leveling to occur.⁵⁹

4.4.2.3 WHY NOT THE 1PL?

The idea that frequency conditions which cells undergo analogical leveling implies that we should draw more general conclusions about which cells in a morphomic distribution are more likely to level. Indeed, given that the 1PL is much more frequent (and less marked) than the 2PL, as described above, we might predict that the 2PL should always level before the 1PL.⁶⁰ Obviously, the Lombard varieties discussed in Section 4.4.1 that show a leveled 1PL directly counter this prediction. However, in the aforementioned case, it is clear that there is a usage-based reason for that specific development—namely the use of the 3SG impersonal for the 1PL. It remains to be seen whether there are cases where the 1PL levels before the 2PL without the external influence of the 3SG impersonal.

There are several reasons for why the 1PL did not level in Piedmontese before the 2PL. Firstly, the 1PL present indicative is homophonous to the corresponding imperative, reinforcing its identity.⁶¹ Secondly, the 1PL ending *-(i)oma* is not identical to any other desinence in the present indicative, unlike the 2PL, which was aligned with its singular counterpart (see the historical explanation above). It is difficult to say whether the only reason that the 1PL did not level in Piedmontese is purely due to its higher frequency or due to a combination of some or all of these other factors.

⁵⁹This is not to say that frequency is the only factor involved in these analogies. Nor does it mean that frequency is the primary driver of analogical change generally. The main takeaway from this section is that frequency in so far as it impacts the acquisition of these sequences directly affects how morphomic distributions are maintained diachronically.

⁶⁰It might also be possible that the higher frequency of the 1PL allows the form to serve as a base from which the 2PL is derived (cf. Albright 2002a).

⁶¹This, of course, is the case in many varieties of Romance (cf. exhortative subjunctives).

4.4.2.4 INTERIM SUMMARY

In the above sections, I have demonstrated that the distribution in Piedmontese, in which the suffix *-sc-* has been introduced to the 2PL, can be explained diachronically in terms of frequency. The 2PL is the least frequent person and number combination cross-linguistically and is thus prone to being influenced by the other cells in the paradigm. That the extension of *-sc-* in Piedmontese was motivated by frequency is evidenced by the fact that the verb ‘to go’ does not show an extension of the *va(d)-* stem, a fact which can be explained by the high frequency of that specific verb. In addition, the 2PL of the imperative does not undergo the extension. Its high frequency allows the N-PATTERN to be preserved in the imperative.

The extension of *-sc-* to the 2PL results in a further narrowing of the domain of the impoverishment rule. We can thus posit the following formalization:

- (15) $[+pres] \rightarrow \emptyset / \text{ ___ } [+2, +pl, +ind]$

1PL IND CONSERVING in Piedmontese

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

This impoverishment rule in (15) does not quite capture the data, specifically the fact that *-sc-* is not present in the 2PL of the imperative. Section 4.6.2 expands on the shortcomings of this rule. For now, however, it is sufficient to note that the general principle of the impoverishment rule accounts for the narrowing.

4.5 LEVELING OF THE SUBJUNCTIVE

This section discusses the analogical leveling of the subjunctive in favor of *-sc-*. As detailed above in Section 4.3.2, there are two types to consider. The first type occurs in several Lombard dialects, e.g. Milanese, in which the subjunctive was leveled after the suffix *-sc-* was introduced to the 1PL of the PRES IND and PRES SUBJ, schematized below in (16). In this type, the leveling in the subjunctive only affected one cell, the 2PL.

(16) 2PL CONSERVING $\Rightarrow$ 2PL IND CONSERVING

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

⇓

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

The second type occurred in dialects found in the area spanning the regions from north-west Italy along the Mediterranean towards Iberia (e.g. Old Occitan, Ligurian, Piedmontese, etc). In this type, the N-PATTERN was subject to leveling of the present subjunctive directly, schematized below:

(17) N-PATTERN $\Rightarrow$ 1/2PL IND CONSERVING

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

⇓

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

As noted in Section 4.3.2, I treat both of these types together and in what follows provide a unified explanation. Section 4.5.1 provides data for the various dialects that level the subjunctive in favor of *-sc-*. The subsequent sections contrast three approaches to the analogy: i) an external analogy, ii) an internal analogy, and iii) morphomic attraction.

4.5.1 DATA FOR LEVELING IN THE SUBJUNCTIVE

4.5.1.1 LOMBARDY

After the extension of *-sc-* to the 1PL, some dialects also leveled the subjunctive in favor of the *-sc-* alternant. To exemplify this, consider the paradigm of the present indicative and present subjunctive of *capì* ‘to understand’ in Milanese (Nicoli, 1983).

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>capíssi</i>	<i>capíisset</i>	<i>capíss</i>	<i>capíssom</i>	<i>capíi</i>	<i>capíssen</i>
PRES SUBJ	<i>capíssa</i>	<i>capíisset</i>	<i>capíssa</i>	<i>capíssom</i>	<i>capíssov</i>	<i>capíssen</i>

Table 4.29: PRES IND/SUBJ of *capì* ‘to understand’ in Milanese (Nicoli 1983: 302)

Here, only the 2PL of the present tense lacks the suffix *-iss-* < Lat. *-isc-*.

The same pattern is also found in the stress pattern of the other conjugations,

e.g. the second conjugation Table 4.30:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>védi</i>	<i>védet</i>	<i>véd</i>	<i>védom</i>	<i>vedíi</i>	<i>véden</i>
PRES SUBJ	<i>véda</i>	<i>védet</i>	<i>véda</i>	<i>védom</i>	<i>védov</i>	<i>véden</i>

Table 4.30: PRES IND/SUBJ of *vedè* ‘to see’ in Milanese (Nicoli 1983: 296)

4.5.1.2 MEDITERRANEAN DISTRIBUTION

Some languages stretching from northwest Italy to the eastern coast of Spain along the coast of the Mediterranean also demonstrate the leveling of the present subjunctive, albeit directly from the N-PATTERN. For example, consider the paradigms from Ligurian (Table 4.31), Upper-Engadinian (Table 4.32), and Old Occitan (Table 4.33) below:⁶²

	1SG	2SG	3SG	1PL	2PL	3PL
IND	<i>patíscu</i>	<i>patísci</i>	<i>patísce</i>	<i>patímu</i>	<i>patì</i>	<i>patísce</i>
SUBJ	<i>patísce</i>	<i>patísci</i>	<i>patísce</i>	<i>patíscimu</i>	<i>patísci</i>	<i>patísce</i>

Table 4.31: PRES IND/SUBJ of *patí* ‘to suffer’ in Ligurian (Azaretti 1977: 220)

	1SG	2SG	3SG	1PL	2PL	3PL
IND	<i>finésch</i>	<i>finéschast</i>	<i>finéscha</i>	<i>finíns</i>	<i>finís</i>	<i>finéschan</i>
SUBJ	<i>finéscha</i>	<i>finéschast</i>	<i>finéscha</i>	<i>finéschans</i>	<i>finéschas</i>	<i>finéschan</i>

Table 4.32: PRES IND/SUBJ of *fin-* ‘to finish’ in Upper-Engadinian (Iliescu & Mourin 1991: 93-94)

In these paradigms, the reflex of Latin *-sc-* is seen throughout the entire present system apart from the 1PL and 2PL of the present indicative. Descriptively, the present subjunctive has been leveled in favor of the *-sc-* alternant.

This pattern is generally replicated elsewhere in these languages, e.g. vowel alterations conditioned by stress in Ligurian (*cöglie* < Lat. *colligere* ‘to gather’):⁶³

⁶²These examples were originally retrieved from Meul (2016).

⁶³However, see Section 4.5.3, where Old Occitan demonstrates a stress pattern which operates independently from the suffix.

	1SG	2SG	3SG	1PL	2PL	3PL
IND	<i>florís(c)</i>	<i>florís</i>	<i>florís</i>	<i>florém</i>	<i>florétz</i>	<i>floríssen</i>
SUBJ	<i>florísca</i>	<i>floríscas</i>	<i>florísca</i>	<i>floriscám</i>	<i>floriscátz</i>	<i>florísca</i>

Table 4.33: PRES IND/SUBJ of *florír* ‘to flourish’ in Old Occitan ([Anglade 1921](#): 282-283)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>cögliu</i>	<i>cögli</i>	<i>cöglie</i>	<i>cüglimu</i>	<i>cügli</i>	<i>cöglie</i>
PRES SUBJ	<i>cöglie</i>	<i>cögli</i>	<i>cöglie</i>	<i>cöglimu</i>	<i>cögli</i>	<i>cöglie</i>

Table 4.34: PRES IND/SUBJ of *cöglie* ‘to pick up’ in Ligurian ([Azaretti 1977](#): 220)

This example illustrates how a vocalic alternation is present in the indicative (*ö* [ø] versus *ü* [y]), but crucially not in the subjunctive, where the root is stressed in both the 1PL and 2PL.

4.5.2 EXTERNAL VERSUS INTERNAL APPROACHES

There are effectively two ways to approach the motivation behind leveling in the present subjunctive. The first constitutes an *external* analogy, in which the subjunctive becomes regularized in analogy with other tenses or moods where the N-PATTERN is not present. The second, on the other hand, assumes an *internal* analogy, in which the subjunctive levels based on the other forms in its local paradigm.⁶⁴ This section argues in favor of the latter approach.

External approaches have been used throughout the literature to account for the leveling of the subjunctive in Lombard (the 2PL IND CONSERVING pattern), although the same logic can be applied to other dialects which level the subjunctive in favor of the *-sc-* alternant. To illustrate this approach, consider the extension of *-sc-* to the 2PL of the present subjunctive in Milanese *capíssov* (compare Standard Italian *capiàte* with no *-isc-*), accompanying a shift of the stress to the root in the other conjugations. [Meyer-Lübke \(1894\)](#) proposes that the 2PL is stressed on the root on the basis of the imperfect

⁶⁴That is, the 1PL and 2PL are leveled based on the singular and 3PL of the subjunctive.

indicative. In these dialects, the imperfect indicative has a columnar accent which occurs on the imperfect marker throughout the paradigm.⁶⁵ Consider the paradigms for the imperfect indicative and present subjunctive in Milanese in Table 4.35.

	SENTÌ ‘to feel’		CAPÌ ‘to understand’	
	IMPF IND	PRES SUBJ	IMPF IND	PRES SUBJ
1SG	<i>sentívi</i>	<i>sénta</i>	<i>capívi</i>	<i>capíssa</i>
2SG	<i>sentívet</i>	<i>séntet</i>	<i>capívet</i>	<i>capíisset</i>
3SG	<i>sentíva</i>	<i>sénta</i>	<i>capíva</i>	<i>capíssa</i>
1PL	<i>sentívom</i>	<i>séntom</i>	<i>capívom</i>	<i>capíssom</i>
2PL	<i>sentívov</i>	<i>sentóv</i> → <i>séntov</i>	<i>capívov</i>	<i>capíi</i> → <i>capíssov</i>
3PL	<i>sentíven</i>	<i>sénten</i>	<i>capíven</i>	<i>capíssen</i>

Table 4.35: IMPERF IND and PRES SUBJ of *sentì* and *capì* in Milanese

Both the regular fourth conjugation verb *sentì* and the *-iss-* verb *capì* have a columnar stress pattern in the imperfect indicative, where the suffix *-ív-* is stressed in both the 1PL and 2PL. Likewise, the stress in the present subjunctive is also columnar. For *sentì*, the stress occurs on the root while for *capì*, the *-íss-* suffix has been extended to the 2PL and is stressed throughout the paradigm. According to Meyer-Lübke (1894), the 2PL of the imperfect indicative exerted analogical pressure on the 2PL of the present subjunctive. In exactly the same way, we might posit a similar comparison between tenses or moods with columnar stress (such as the imperfect indicative or imperfect subjunctive) to the present subjunctive in Ligurian, Upper-Engadinian, and Old Occitan, which all show a leveling in the subjunctive but the N-PATTERN in the present indicative.

The question arises of why the imperfect indicative would exert analogical pressure on the present subjunctive. From a syntactico-semantic perspective, neither of these constructions share morphosyntactic features. Accordingly, the forms in these two tenses are not any more similar than any other tenses or moods. The imperfect indicative is constructed by the addition of the according theme vowel + *-v-*, while the present

⁶⁵That is, the accent appears regularly on the second syllable throughout the paradigm.

subjunctive consists of the root with the personal endings.⁶⁶ Thus, from a formal perspective, the choice of the imperfect indicative as the model for the analogy is *ad hoc*. Furthermore, the very fact that we can easily extend this analogy to the present indicative as well demonstrates the weakness of this explanation. While it may have been the case that learners associated the present subjunctive with the imperfect indicative, there is simply no evidence that supports this hypothesis.

The alternative, an *internal* analogy, assumes that the subjunctive was leveled purely based on forms within the present subjunctive.⁶⁷ Thus, the *-sc-* was extended to the 1PL and 2PL on the basis of the singular and 3PL forms, where the suffix *-sc-* was present. In this sense, the analogy can be framed in terms of learners attempting to regularize the marking of the subjunctive, along the lines of a natural morphosyntactic class.

This approach can be easily reconciled with other linguistic and cognitive factors. For example, the maintenance of the pattern in the indicative as opposed to the subjunctive can be explained in terms of frequency or markedness.⁶⁸ For the 2PL CONSERVING pattern in particular, the maintenance of such a pattern relies on learners having strong evidence for the 2PL of the subjunctive. As noted in Section 4.4.2, this cell is the most marked and least frequent cell of the paradigm, suggesting that it would be difficult for learners to maintain its irregularity. This consideration makes the change from (18-a) to (18-b) almost trivial.

⁶⁶ As well as any other relevant present tense marking, like the theme vowels or *-sc-* suffix.

⁶⁷ This approach is not new, cf. Mourin (1977).

⁶⁸ That the subjunctive is more marked than the indicative is supported both in terms of the lower relative frequency of the subjunctive, as well as later acquisition of the subjunctive by both L1 learners (Clark 1985 for Romance) and L2 learners (Gragera 2000, Kaufman 2011, *a.o.* for Spanish).

Additionally, subjunctives are generally marked overtly with more linguistic material, e.g. the Modern Greek subjunctive (Philippaki-Warbuton 1993: 301).

- (18) a. 2PL CONSERVING in Lombard (Cremonese, Bergamasque, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

- b. 2PL IND CONSERVING in Lombard (Milanese, Bresciano, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

An internal approach for the intrusion of *-sc-* into the 1PL and 2PL of the subjunctive can also be framed in terms of the regularization of the stem in the present subjunctive, based on the model provided by the singular and 3PL. If we put aside stress for the moment, we may consider that for most verbs, the stem used for the 1PL and 2PL of the subjunctive is identical for all person and numbers of the present subjunctive. Thus, a regular first conjugation verb in Old Occitan shares the same general stem (*cant-*) for all person and numbers of the present indicative and subjunctive, seen below in Table 4.36:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>cant</i>	<i>cántas</i>	<i>canta</i>	<i>cantám</i>	<i>cantátz</i>	<i>cántan</i>
PRES SUBJ	<i>cant</i>	<i>cantz</i>	<i>cant</i>	<i>cantém</i>	<i>cantétz</i>	<i>cánten</i>

Table 4.36: PRES IND/SUBJ of *cantár* ‘to sing’ in Old Occitan (Anglade 1921: 269-276)

The internal coherence of the stem in the subjunctive could thus be applied to forms with *-sc-*, producing forms with *-sc-* in the 1PL and 2PL:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>florís(c)</i>	<i>florís</i>	<i>florís</i>	<i>florém</i>	<i>florétz</i>	<i>floríssen</i>
PRES SUBJ	<i>florísca</i>	<i>floríscas</i>	<i>florísca</i>	<i>floriscám</i>	<i>floriscátz</i>	<i>floríscan</i>

Table 4.37: PRES IND/SUBJ of *florír* ‘to flourish’ in Old Occitan (Anglade 1921: 282-283)

If we now take stress into account, we can see that Old Occitan maintained the

older stress pattern (in the shape of the N-PATTERN) for both the *-a-* and *-sc-* classes, while introducing *-sc-* to the 1PL and 2PL. This separation of stress and stem suggests that speakers extended *-sc-* in order to mark the present tense, rather than to produce a columnar stress pattern based on a comparison of the subjunctive to external tenses or moods, such as the imperfect indicative, which had columnar stress.

This explanation is heavily reminiscent of accounts for analogical leveling that assume that learners make generalizations based on ‘informative’ bases (cf. Albright 2002a, 2005). If, for example, the 3SG PRES SUBJ of *cantár* ‘to sing’ was the most informative (in Old Occitan), then the production of the 1PL and 2PL of the present subjunctive of *florír* ‘to flourish’ as *floriscám* and *floriscátz* would follow from 3SG *florísca*.⁶⁹ Regardless of which stem was the base, it is clear that it is easy to motivate an internal leveling of the subjunctive.

4.5.3 MORPHOMIC COHERENCE?

Further evidence for an internal analogy can be deduced by comparing the internal approach with a morphomic approach. This section focuses on the approach set forth in Maiden (2003, 2012, 2018), which describes the leveling in the subjunctive as the intersection between morphomic patterns.⁷⁰ While this account was originally proposed to account for the 1/2PL IND CONSERVING distribution, the account is equally applicable to the leveling in the subjunctive in the dialects found in Lombard, which have also introduced *-sc-* to the 1PL.

In order to account for the leveled forms in the present subjunctive, Maiden (2003, 2012, 2018) has proposed that the leveling was due to a conflation between two existing

⁶⁹This is not to say that the 3SG is in fact the correct base upon which the rest of the subjunctive was formed. It is merely a hypothetical proposal to illustrate how *-sc-* can be extended on the basis of informativity.

⁷⁰Such an approach sits in between an external and internal analogy.

morphomic patterns, the aforementioned N-PATTERN and the L-PATTERN. In short, the L-PATTERN refers to a correspondence between the 1SG of the present indicative with the entire present subjunctive.⁷¹ Thus, in many Romance languages, we find paradigmatic alignments as follows:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i> digo </i>	<i> dices </i>	<i> dice </i>	<i> decimos </i>	<i> decís </i>	<i> dicen </i>
PRES SUBJ	<i> diga </i>	<i> digas </i>	<i> diga </i>	<i> digamos </i>	<i> digáis </i>	<i> digan </i>

Table 4.38: PRES IND/SUBJ of *decir* ‘to say’ in Spanish (Maiden 2003: 45)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i> vail </i>	<i> vaus </i>	<i> vaut </i>	<i> valons </i>	<i> valez </i>	<i> valent </i>
PRES SUBJ	<i> vaille </i>	<i> vailles </i>	<i> vaille </i>	<i> vailliens </i>	<i> vailliez </i>	<i> vaillent </i>

Table 4.39: PRES IND/SUBJ of *valier* ‘to be worth’ in Old French (Maiden 2003: 46)

In these examples, the present subjunctive aligns with the 1SG of the present indicative in terms of root shape. In Table 4.38, the stem *dig-*, found in the subjunctive and 1SG of the indicative, alternates with *dic-* elsewhere. The same distribution is seen in Table 4.39, with *vail-* alternating with *va(l)-*.

As with the N-PATTERN, the origin of this pattern has a clear phonological cause (cf. Maiden 2018: 84-90). The environment of a consonant followed by an unstressed prevocalic *i* and *e* underwent many changes in different Romance varieties, including palatalization and affrication.⁷² This change largely affected second and fourth conjugation verbs, where the root in the present subjunctive and 1SG of the present indicative was followed by one of these vowels. The palatalization in these forms created the distinction seen in Table 4.39, where the *vail-* stem was derived from **vali-* < Lat. *vale-* ‘be worth’ before vowels.⁷³

⁷¹The term L-MORPHOME, which I refer to here as L-PATTERN, can be derived from its paradigmatic shape when presented as in Table 4.38 and Table 4.39.

⁷²This change is referred to as ‘the yod effect’ by Maiden (2018), as prevocalic *i* and *e* became [j], causing the various changes that we see.

⁷³The 3PL ending complicates this picture. In many Romance varieties, the ending *-iunt* which would

This pattern was reinforced by other phonologically-induced changes that created the same distribution, specifically palatalization of [k] and [g] before front vowels.⁷⁴ For example, in Table 4.38, both the 1SG of the present indicative and the present subjunctive of Lat. *dīcō* were not palatalized, as the stem *dīc-* was not followed by a front vowel: Lat. 1SG PRES IND *dīcō* ‘I say’ and PRES SUBJ *dīcam*, *dīcās*, etc. In Spanish, the [k] was voiced intervocalically, producing *digo*, *diga*, etc.⁷⁵ However, the other forms of the present indicative, apart from the 3PL, were following by a front vowel. The consonant preceding the front vowel in these forms was susceptible to palatalization, creating the alternation seen in Table 4.38: Lat. *dīcis* > Sp. *dices*.⁷⁶

To summarize the phonological conditioning of the morpheme, the L-PATTERN was created by regular sound changes that created palatalized stem variants along the lines seen above in Table 4.38 and Table 4.39. For the former, the palatalized forms are found in the present indicative (but crucially not the 1SG), while for the latter, the palatalized alternants are found throughout the subjunctive as well as the 1SG of the present indicative.

This pattern has been extended analogically to other verbs in which we would not regularly expect a palatalized outcome, confirming its status as a morpheme. For example, we see extension of L-PATTERN alternants to the 1SG in the descendants of Lat. *possum* ‘I am able’ (Maiden 2018: 94):

have caused palatalization was replaced by *-ent*, producing the L-PATTERN, as the 3PL indicative does not have a palatalized form. However, some varieties in Italo-Romance generalized the *-iunt* ending, leading to the creation of a so-called U-PATTERN (Maiden 2018: 85). For this reason, Maiden refers to the morpheme as a whole as the L/U-MORPHOME.

⁷⁴Notably this change is distinct from the aforementioned palatalization.

⁷⁵Mod. Sp. [diyo], [diya], [diɣa], etc.

⁷⁶The 3PL ending *-unt* was replaced by the ending *-ent*, creating a form that was susceptible to palatalization (cf. fn. 73).

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>posso</i>	<i>podes</i>	<i>pode</i>	<i>podemos</i>	<i>podeis</i>	<i>podem</i>
PRES SUBJ	<i>possa</i>	<i>possas</i>	<i>possa</i>	<i>possamos</i>	<i>possais</i>	<i>possam</i>

Table 4.40: PRES IND/SUBJ of *poder* ‘to be able’ in Portuguese (Maiden 2018: 94)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>pós</i>	<i>pó</i>	<i>pó</i>	<i>pudómma</i>	<i>pudì</i>	<i>pón</i>
PRES SUBJ	<i>pósa</i>	<i>pós</i>	<i>pósa</i>	<i>pósen</i>	<i>pósi</i>	<i>pósen</i>

Table 4.41: PRES IND/SUBJ of *pudài* ‘to be able’ in Cascinagrossa (Castellani 2002: 91)⁷⁸

In both Table 4.40 and Table 4.41, the L-PATTERN has seemingly exerted an analogical influence on *possum*. The key takeaway of these forms is that in addition to the N-PATTERN (and its derivatives) found throughout the verbal system, there was also an additional pattern in which the entire subjunctive followed a singular pattern of morphological expression, which is accompanied by the 1SG present indicative.

In the view of Maiden (2003, 2018) and most crucially Maiden (2012), the languages which extended the *-sc-* suffix to the 1PL and 2PL of the present subjunctive did so because the L-PATTERN attracted the forms with *-sc-*. In this sense, the leveling of the subjunctive is the result of the intersection of the L-PATTERN and the N-PATTERN where the former simply overrode the latter in the subjunctive due to the identity of the 1SG PRES IND. Maiden proposes that this change occurred by means of the 1SG present indicative extending its form to the subjunctive. Thus, the analogical pressure for the leveling was not just the singular and 3PL of the subjunctive but also the 1SG of the present indicative.

Maiden evidences his claim by virtue of how this change is often accompanied by a shift in the stress pattern in the 1PL and 2PL from the ending to the suffix, seemingly because the stress in the 1SG present indicative is on the suffix *-sc-* (Maiden 2003: 48).⁷⁹

⁷⁸This is a dialect of Piedmontese which does not show syncretism of the 2SG and 2PL.

⁷⁹In other verb conjugations which lack the suffix *-sc-*, the stress is moved to the root. Expectedly, we find vowel alternations spreading to the 1PL and 2PL of the subjunctive (Maiden 2003: 47-48).

To summarize this approach, there exists a morphomic pattern, the L-PATTERN which exists alongside the N-PATTERN. As the pattern consists of the present subjunctive and the 1SG, it can theoretically serve as an attractor for the 1PL and 2PL of the present subjunctive, allowing those forms to assimilate to the other cells in the L-PATTERN. Fundamentally, the L-PATTERN is overriding the N-PATTERN in the subjunctive.

4.5.3.1 INTERNAL ANALOGY VERSUS MORPHOMIC ATTRACTION

The two approaches differ in the following way. An internal approach posits that the 1PL and 2PL are leveled under pressure of the rest of the present subjunctive paradigm, on the assumption that learners want to have consistent mapping across these forms. The morphomic approach suggested by Maiden differs only slightly in that learners map the 1PL and 2PL to a template that additionally refers to the 1SG of the present indicative. Thus, the only real distinction is whether learners feel the need to reference the 1SG of the present indicative. Without reference to the 1SG, it is impossible to determine which approach is correct for a given leveled paradigm.

I argue that it is more likely that speakers leveled the 1PL and 2PL of these paradigms (especially those with *-sc-*) in terms of the subjunctive rather than the L-PATTERN, on the basis that i) there is little evidence that forms with *-sc-* in the dialects of interest reference an L-PATTERN, and ii) the patterns of leveling in the paradigm of ‘to go’ strongly indicate that the development is internal and do not reference the L-PATTERN. This is not to say that the L-PATTERN never constitutes a meaningful analogical attractor; it simply does not provide any explanatory power in the cases of the leveling of *-sc-*.

Beginning with the first argument, we can note that there are three distributions described in this chapter that have a leveled subjunctive, shown below in (19):

(19) Distributions with Leveled Subjunctive

- a. 2PL IND CONSERVING in Lombard (Milanese, Bresciano, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

- b. 1/2PL IND CONSERVING in Ligurian, Old Occitan, etc.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

- c. 1PL IND CONSERVING in Piedmontese

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

It is true that each of these distributions contains the cells that participate in the L-PATTERN. However, it is also the case that several cells have *-sc-* which are not in the L-PATTERN (e.g. the 2/3SG). There is simply no reason that, out of the cells in the present indicative, the 1SG would exert any specific influence.

There is also no overt alignment of these distributions with the L-PATTERN. If the L-PATTERN exerted such a strong influence that it was able to level the 1PL and 2PL subjunctive, why did it not also remove *-sc-* from the cells in the present indicative that do not participate in the L-PATTERN? To my knowledge, no Romance variety demonstrates a distribution of *-sc-* where the suffix is limited to the L-PATTERN. Such a distribution would provide certain evidence that the L-PATTERN is the cause of the leveling in the subjunctive; however, the absence of such evidence suggests that the change in *-sc-* is

simply orthogonal to other examples of analogy involving the L-PATTERN.⁸⁰

The second argument focuses on the distribution of suppletive forms in the verb ‘to go’ in the dialects that show leveling in the subjunctive. While several do not undergo leveling in this paradigm (see Section 4.4.1.3), those that do are separated into two types. The first type shows the *va(d)*- stem leveled throughout the subjunctive, as can be seen below in 4.42 and 4.43:⁸¹

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>vad</i>	<i>vas</i>	<i>va</i>	<i>andoma</i>	<i>andeve</i>	<i>van</i>
PRES SUBJ	<i>vada</i>	<i>vade</i>	<i>vada</i>	<i>vado</i>	<i>vade</i>	<i>vado</i>

Table 4.42: PRES IND/SUBJ of *andé* ‘to go’ in Piedmontese (Griva 1980: 84)

⁸⁰A small point that reinforces this argument is the fact that stress still exhibits the N-PATTERN distribution in Old Occitan (see Table 4.37) in the 1PL and 2PL. If the 1SG of the present indicative played an important role in the leveling of these forms, we would expect stress on the suffix, rather than the descence.

⁸¹This example does provide evidence for an L-PATTERN, as the stem *vad-* occurs exactly in the distribution of the pattern. However, it also demonstrates a fundamental issue with morphomic categorization. If we analyze *vad-* as one stem, how should we classify the other stems in the present indicative? Obviously the *and-* stems indicate a preservation of the N-PATTERN, but it is unclear how the *va-* stem found in the 2/3SG and 3PL should fit in. If we take all the *va(d)*- forms together, we miss the generalization that *vad-* occurs not only in the subjunctive, but also the 1SG. If we separate them, we lose the N-PATTERN generalization remaining in the present indicative.

This issue can be exemplified by the following. Maiden (2012) classifies the following pattern from a variety of Corsican as an example of the L-PATTERN:

- (i) ‘to go’ in Corsican variant (Maiden 2012: 50)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>'vogu</i>	<i>'vai</i>	<i>'va</i>	<i>an'demu</i>	<i>an'deti</i>	<i>'vani</i>
PRES SUBJ	<i>'vogi</i>	<i>'vogi</i>	<i>'vogi</i>	<i>'vogimi</i>	<i>'vogiti</i>	<i>'vogini</i>

However, in the same article, Maiden (2012) proposes that the same verb in Italian shows the N-PATTERN:

- (ii) *andàre* ‘to go’ in Italian (Maiden 2012: 41)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>vàdo</i>	<i>vài</i>	<i>va</i>	<i>andiàmo</i>	<i>andiàte</i>	<i>vànno</i>
PRES SUBJ	<i>vàda</i>	<i>vàda</i>	<i>vàda</i>	<i>andiàmo</i>	<i>andiàte</i>	<i>vàdano</i>

It is certainly peculiar that the *vad-* stem in Italian has no special status, but the *vag-* stem in the Corsican variant does. The L-PATTERN is thus only referenced by Maiden in order to explain the leveling in the 1PL and 2PL of the subjunctive. The fact that *vad-* occurs in Italian in the subjunctive (besides the 1PL and 2PL) and the 1SG present indicative is simply a coincidence according to this view.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	vagu	vai	va	<i>andamu</i>	<i>andé</i>	van
PRES SUBJ	vaghe	vaghi	vaghe	vaghimu	vaghi	vaghe

Table 4.43: PRES IND/SUBJ of *andà* ‘to go’ in Ligurian (Azaretti 1977: 208)

However, in some varieties, it is the stem in the 1PL and 2PL of the subjunctive that is generalized to the entire subjunctive, without any reference to the indicative. For example, in Old Occitan, the verb *anar* ‘to go’ also exhibits a completely leveled present subjunctive.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	vau/vaux	vas	va, vai	<i>anám</i>	<i>anátz</i>	van
PRES SUBJ	<i>an</i>	<i>ans</i>	<i>an</i>	<i>aném</i>	<i>anétz</i>	<i>anon</i>

Table 4.44: PRES IND/SUBJ of *anar* ‘to go’ in Old Occitan (Anglade 1921: 278-279)

There is evidence for N-PATTERN variants in the subjunctive: e.g. *vaza*, *vazas*, *vaza*, and *vazan*. Crucially, there are no such variants for the 1PL and 2PL subjunctive. Notably, its stress pattern still maintains the N-PATTERN distribution, despite the leveling of the stem, mirroring the pattern seen in the *-sc-* verbs (see Table 4.37).⁸²

Another example can be found with the verb ‘to go’ in Upper-Engadinian:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	vegn	vest	vo	<i>giains</i>	<i>giais</i>	vaun
PRES SUBJ	<i>giaja</i>	<i>giajast</i>	<i>giaja</i>	<i>giajans</i>	<i>giajas</i>	<i>giajan</i>

Table 4.45: PRES IND/SUBJ of *ir* ‘to go’ in Upper-Engadinian (Scheitlin, 1962)

For Upper-Engadinian, the *gia-* stem [ja/ε-], which is originally from Latin *īre* ‘to go’, has been generalized from the 1PL and 2PL of the present subjunctive to the rest of

⁸²An analysis involving impoverishment is also able to account for the distribution of *-sc-* and ‘to go’ in Old Occitan, where *-sc-* has been extended to the 1PL and 2PL of the subjunctive, but the *an-* stem (as opposed to the *va-* stem) has been extended to the rest of the subjunctive. For both the forms with *-sc-* and the verb *anar* ‘to go’, the 1PL and 2PL of the present indicative are impoverished. The difference between *florír* and *anar* is that *-isc-* is conditioned by the present tense, while the *va-* stem is conditioned by both tense and mood. Both verbs follow the newly formed 1/2PL IND CONSERVING template, as the alternants are limited to the present indicative. However, as *-sc-* and the stem *an-* express different features, their distributions are able to differ outside of the purview of the N-PATTERN.

the paradigm, the exact opposite leveling as that found in Piedmontese and Ligurian, where the *va(d/g)*- stem has been extended to the 1PL and 2PL of the subjunctive.⁸³

These patterns directly contradict the L-PATTERN. A leveling in accordance with the L-PATTERN would either result in subjunctive forms that all have the stem *va-* or a 1SG present indicative with the *an-/gia-* stem.⁸⁴

The alternative explanation assumes that the present subjunctive was leveled based on paradigm-internal forms, simply choosing the stem found in the 1PL and 2PL, rather than the stem in the other persons. It seems to me that Maiden must accept such a view (or something similar) in order to account for these distributions. If we have to assume that some paradigms in the language level the subjunctive on an internal basis, it is more economical to assume that all occurrences of leveling in the subjunctives in favor of an alternant in those dialects occurred on an internal basis, especially those which do not otherwise show an L-PATTERN distribution with forms suffixed with *-sc-*.

⁸³A further example of a leveled subjunctive in the paradigm of ‘to go’ can be found in Bergamasque, which otherwise shows a 2PL CONSERVING pattern:

(i) PRES IND/SUBJ of *andàa* ‘to go’ in Bergamasque (Zanetti 2005: 172)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>'ndó</i>	<i>'ndé</i>	<i>và</i>	<i>và</i>	<i>andì</i>	<i>và</i>
PRES SUBJ	<i>'ndaghe</i>	<i>'ndaghet</i>	<i>indaghe</i>	<i>indàghe</i>	<i>indàghev</i>	<i>'ndaghe</i>

In this latter case, the N-PATTERN has been replaced by reflexes of *'nd-* < *and-* in the subjunctive and in the 1SG and 2SG of the present indicative. The fact that the subjunctive has been leveled here is quite surprising, as the subjunctive has not undergone leveling elsewhere in other verbs. It is clear from the distribution of *va-* and *'nd-* stems in the indicative that a specialized account is required to explain this distribution, which I leave for future work.

⁸⁴One might make the argument that these paradigms never inherited the N-PATTERN to begin with and thus do not present a counterexample to the expected paradigms. In other words, the distributions in Old Occitan and Upper-Engadinian never had the N-PATTERN to begin with. However, given the location of the dialects in question, as well as the rest of the verbal system of said dialects, this seems unlikely. Rather, I assume that these patterns represent historical distributions of the N-PATTERN that have since been leveled in certain ways.

The same argument is made by Maiden (2018: 200): ‘...what we are observing is not peculiar to the verb ‘go’, but rather a manifestation of the more general tendency...for all root allomorphy to be *eliminated* from the present subjunctive over wide areas of the România [i.e. the Romance-speaking area].’

Furthermore, I suggest that the reason why the 1PL and 2PL form of ‘to go’ was extended rather than the other forms may have something to do with the high frequency of those forms. If, for example, the 1PL of the subjunctive is used as an exhortative, as we have seen in other varieties, it may motivate the choice in stem. Other less frequent verbs may not have had such a frequent 1PL form, allowing the stem in the singular and 3PL to replace the 1PL and 2PL. Admittedly, it is difficult to show that this was indeed the case, given the general dearth of information regarding this point, and thus I leave this for further investigation. However, it does make a verifiable prediction that may explain the choice in stem.

4.5.4 SUMMARY

In the last few sections, I have discussed the process by which *-sc-* was extended to the 1PL and 2PL of the present subjunctive. The dialects described all inherited the N-PATTERN; however, this distribution has been obscured in the subjunctive.

I have argued that the leveling of *-sc-* throughout the present subjunctive is best explained via an internal analogy, in which learners are biased in favor of morphological coherence. That is, learners prefer to learn patterns which conform to morphological classes. In doing so, I have also challenged two other explanations, namely that i) *-sc-* was introduced based on an external analogy, or ii) morphomic attraction to the L-PATTERN caused *-sc-* to be introduced to the 1PL and 2PL.

An external approach is hard to motivate, given that it is not clear what the present subjunctive should be compared to. An analogy with the imperfect indicative, as proposed by Meyer-Lübke (1894), is unpersuasive, given the *ad hoc* nature of the comparison.

Likewise, I have argued that the morphomic approach posited by Maiden (2003,

2012, 2018) is not satisfactory. Morphomic attraction to the L-PATTERN would be convincing if the present indicative was restructured such that only the 1SG PRES IND was associated with the *-sc-* suffix. Additionally, the verb ‘to go’ lends itself to an internal explanation, rather than a morphomic one, as the present subjunctive is occasionally leveled in favor of the alternant in the 1PL and 2PL. This result is not accounted for via morphomic attraction. In this sense, the L-PATTERN is not informative in terms of the development of *-sc-*, despite having descriptive power with regards to other analogical changes.

FORMALIZATION This section briefly provides a formalization for the distributions discussed above.⁸⁵

First, the distribution in which the subjunctive has been leveled in favor of *-sc-* after *-sc-* was extended to the 1PL can be described as follows:⁸⁶

(20) [+pres] → ∅ / ____ [+2, +pl, +ind]

2PL IND CONSERVING in Lombard (Milanese, Bresciano, etc.)

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

A similar impoverishment rule can be posited for the distribution in which only the subjunctive has undergone leveling:

⁸⁵As a reminder, the purpose of these formalizations is not to prove that speakers encoded the changes in this way. Rather, they demonstrate that it is possible to encode the morphomic distributions without making reference to a morphomic entity in the grammar. In this sense, I hope to show that a morphome is synchronically epiphenomenal.

⁸⁶Here as elsewhere, I only refer to the specific features in the bundles generally. It may be the case that slightly altered features are necessary in order to account for the nuances of the syntactic structure of the verbal systems in these varieties.

(21) [+pres] → ∅ / ____ [+part, +pl, +ind]

1/2PL IND CONSERVING in Ligurian, Old Occitan, etc.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND						
PRES SUBJ						

4.6 CONCLUSION

In this chapter, I have investigated two pathways affecting the N-PATTERN in Romance. In both, the domain of the suffix *-sc-* has been extended so as to eliminate the alternant *-∅-* over time. This gradual change has resulted in various dialects reflecting different stages of development, allowing for a thorough discussion of the reasons for change in each case.

In order to motivate the analogies that have taken place, I have identified two biases affecting learners, restated below:

- Usage and frequency determine the maintenance of morphomic distributions.
- Learners are biased to change morphomic distributions to fit with natural classes expressed with morphosyntactic features.

These biases are evidenced directly in the types of change seen. The introduction of *-sc-* to specific person and number combinations can be motivated in terms of usage, in so far as that usage affects the frequency of the relevant forms. The subsequent changes in the input that a learner receives motivates the introduction of default morphology, allowing for the leveling out of the *-∅-* alternant.

The leveling of the present subjunctive in favor of the *-sc-* variant can be based on the desire to unify a more marked morphological category, in terms of a cohesive natural

class. Other explanations have been proposed; however, these struggle to account for various aspects of the development.

The rest of this conclusion is dedicated to addressing two minor points. The first pertains to the introduction of *-sc-* to other moods, besides the indicative, subjunctive, and imperative. The second provides a more in-depth synchronic analysis of the distribution in Piedmontese (see Section 4.4.2.1). Here, I hope to provide further evidence that we can account for morphomic distributions via impoverishment rules rather than morphological templates.

4.6.1 INTRODUCTION OF *-SC-* TO THE FUTURE AND CONDITIONAL

There does exist some data in the dialects that puts into question the generalization that speakers want to contrast the present tense with the other tenses; however, in this section, I argue that it actually serves to reinforce the assumption that speakers prefer to assign morphological marking to natural morphosyntactic categories.

Some Lombard dialects variously reform their future and conditional moods to conform with the present tense. For example, the reflex of *-sc-* in Milanese, *-iss-*, is found throughout the future and conditional, while the present tense otherwise shows a 2PL IND CONSERVING pattern (cf. (10-c)). Thus, we find the following forms for the future and conditional in Milanese (Nicoli 1983: 302-303):

	1SG	2SG	3SG	1PL	2PL	3PL
FUT	<i>capissaróo</i>	<i>capissarée</i>	<i>capissará</i>	<i>capissarémm</i>	<i>capissaríi</i>	<i>capissaránn</i>
COND	<i>capissaría</i>	<i>capissariét</i>	<i>capissaría</i>	<i>capissaríom</i>	<i>capissaríov</i>	<i>capissarién</i>

Table 4.46: FUT IND and PRES COND of *capì* ‘to understand’ in Milanese

This is not to say that the intrusion of the suffix *-sc-* in either the future and conditional conditions the leveling in the present tense. For example, Paves, which in the present tense shows a 2PL CONSERVING pattern, lacks the suffix in the conditional,

despite leveling the 1PL in the present tense (cf. Table 4.19).

	1SG	2SG	3SG	1PL	2PL	3PL
COND	<i>finärisi</i>	<i>finärisät</i>	<i>finärisä</i>	<i>finärisäm</i>	<i>finärisäv</i>	<i>finärisän</i>

Table 4.47: PRES COND of *finì* ‘to finish’ in Paves (Comoletti 2005: 71)

The *-is-* in the forms in Table 4.47 is not to be confused with the reflex of *-sc-* as discussed throughout this chapter. Rather, these forms are the result of the addition of the imperfect subjunctive marker to the conditional. Thus, the conditional, itself from the periphrastic combination of the infinitive with the imperfect indicative of Lat. *habēre* ‘to have’, has an appended *-is-*. This *-is-* is itself from the pluperfect subjunctive *-iss-* in Latin and is consequently not from the suffix *-sc-*. The forms in Table 4.47 contrast with the conditional in Milanese (Table 4.46), where *-iss* < Lat. *-sc-* has intruded between the root *cap-* and the marker for the conditional *-rì-*.

Distributions where *-sc-* has intruded upon the future/conditional challenge the notion that *-sc-* denotes the present tense. However, I argue that these cases are still amenable to the general observation that the morphological material is associated with morphosyntactic distributions. Both the future and the conditional, despite not being present tense *per se*, can be grouped with the present as distinctly ‘non-past’. There is thus a difference between the languages where *-sc-* is restricted to the strictly present tense forms and those that restrict the suffix to non-past tense forms. However, this difference does not prove to be crucial for our purposes; both distributions present a situation in which some set of forms is distinguished from an ‘elsewhere case’ by a set of morphological exponents (associated either with present or non-past). In terms of the leveling that affects the present tense, it is immaterial whether the natural class reflected by *-sc-* is the present tense feature or a non-past feature.

Maiden (2018: 243) explicitly argues against an analysis associating the *-sc-* with ‘non-past’ as it would also affect the 1PL and 2PL of the present indicative and sub-

junctive. Since these cells do not show the expected ‘non-past’ morphology, a marker like *-sc-* simply cannot mark a non-past feature (or a present feature for that matter). This criticism is largely addressed by the points made in Section 4.2.4 when discussing impoverishment.⁸⁷

To summarize, the alternations in the present tense in Romance obscure a deeper opposition between the present (or non-past) tense forms and those of the rest of the paradigm. Recognizing this fact, we can explain the tendency to level the present tense in favor of the morphological variants that only appear in the present as a diachronic reinforcement of this opposition. Additionally, from a synchronic perspective, we can associate alternants in the 1PL and 2PL with the elsewhere forms found in other (non-local) paradigms of a given verb. As for the future and conditional forms, these can be explained as the reanalysis of the present tense marking as non-past marking.

4.6.2 TEMPLATE VERSUS RULES: PIEDMONTESE

This section discusses the distribution in Piedmontese and how the synchronic analysis presented in Section 4.2.4 provides a potential insight into how we can disambiguate a ‘rules-based’ analysis like that in Embick (2016) from a templatic approach as that adopted by Maiden (2018).

To reiterate, the distribution of *-sc-* in Piedmont is particularly idiosyncratic. The relevant present tense forms are shown below:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>finisso</i>	<i>finisse</i>	<i>finís</i>	<i>finioma</i>	<i>finisse</i>	<i>finisso</i>
PRES SUBJ	<i>finissa</i>	<i>finisse</i>	<i>finissa</i>	<i>finisso</i>	<i>finisse</i>	<i>finisso</i>
PRES IMP	–	<i>finiss</i>	<i>finissa</i>	<i>finioma</i>	<i>finì</i>	<i>finisso</i>

Table 4.48: PRES IND/SUBJ of *finí* ‘to finish’ in Piedmontese (Grosso 2002: 96-99)

⁸⁷The proposal in Maiden (2018: 244-245) as involving the specific Italian lexeme *rete* ‘net’ being phonologically similar to the 2PL FUT suffix *-rete* is striking and seems highly implausible to me.

Crucially, while the forms in the indicative and subjunctive are leveled in the 2PL, this is not the case in the imperative, where they retained the old imperative form. An additional relevant detail is the fact that the 2PL of the present indicative and subjunctive are identical to their singular counterparts.

In order to account for this distribution synchronically, a templatic approach simply assumes that a new morpheme has been created following the bolded forms in Table 4.48. The fact that the 2PL of the imperative does not coincide with the indicative and subjunctive even reinforces the notion that the morpheme is idiosyncratic in terms of its paradigmatic configuration.

For a rules-based analysis, we could simply add an additional disjunct environment to the impoverishment rule listed in (15), specifically targeting the 2PL imperative. This modification is represented as follows:⁸⁸

$$(22) \quad [+pres] \rightarrow \emptyset / \text{ ___ } [+2, +pl, +ind], [+2, +pl, +imp]$$

As it stands the rule in (22) and a templatic approach are effectively equivalent.⁸⁹ Both account for the presence of *-sc-* in the relevant cells. However, both also fail to account for the generalization that the 2PL is identical to the 2SG.

I suggest that an additional impoverishment rule targeting the 2PL more generally may allow for a unified analysis. The crucial observation is that the homophony of the 2PL and 2SG can be characterised as a meta-syncretism (Williams 1994, Bobaljik 2002, Harley 2008, Kramer 2016), which in DM may be captured by an impoverishment rule. In this case, the relevant feature that must be impoverished is the [+pl] feature,

⁸⁸In here and what follows, I assume [+imp] stands for whatever features are required to account for the imperative. A full account may require more specificity.

⁸⁹However, this impoverishment rule differs in that it introduces a new condition. The formalization of the other changes have involved reducing the context of the elsewhere case. In this sense, the impoverishment rule in (22) actually introduces complexity, which may be unexpected. This issue will be resolved below.

effectively allowing 2SG forms to be spelled out for 2PL forms.

(23) $[+pl] \rightarrow \emptyset / \text{ ___ } [+2, -imp]$

This rule allows the form *finisse* to be the morphological exponent of both the 2SG and the 2PL.

As for the original formulation of the impoverishment rule targeting the [-past] node, this can be restated with the original N-PATTERN formulation. The two rules can be seen below:

(24) Piedmont Impoverishment rules

- a. $[+pl] \rightarrow \emptyset / \text{ ___ } [+2, -imp]$
- b. $[+pres] \rightarrow \emptyset / \text{ ___ } [+part, +pl]$

If ordered first, (24-a) bleeds (24-b), allowing for the presence of *-sc-* in the 2PL.⁹⁰ The advantage of this approach over the alternatives is that it unifies the morphomic distribution as well as the syncretism under a single mechanism, supporting the notion that speakers can use such rules to account for morphological idiosyncrasies.⁹¹

In sum, it is possible that the distribution of *-sc-* in Piedmont shows evidence that the morphomic distribution is synchronically stored in terms of rules, rather than as a template. Further, it suggests that if we consider whether morphomic templates consist of a set of rules, we may be able to derive deeper explanations for the surface distributions that we see.

⁹⁰ Admittedly, this account requires rule ordering to apply to impoverishment rules. This is a significant drawback, and I will not provide a full justification here. However, I suggest that this is something that can be examined in future work.

⁹¹ Likewise, this approach may also be used to explain the inconsistent distributions seen in the verb ‘to go’ as discussed at various points in this chapter. For example, in Piedmontese, the verb ‘to go’ does not show leveling in the 2PL in the present indicative (similar to the present imperative). With the correct tweaking of conditioning environments, these idiosyncrasies can be explained via impoverishment rules as suggested here.

5

Conclusion

5.1 OVERVIEW

This dissertation has investigated how two types of analogy, proportional change and paradigmatic leveling, interact with *non-root intra-paradigmatic alternations*. These alternations come in many types and distributions, but most insightful are irregular (or morphomic) distributions, as these are difficult to explain from both diachronic and synchronic perspectives. In examining some of these distributions in depth, I have identified several biases which affect how these distributions change over time. Addi-

tionally, the diachronic changes discussed in this dissertation have provided valuable insight into the synchronic status of irregular distributions, with a specific emphasis on their learnability as well as their extendability to new verbs. The next section outlines the contributions in greater detail, before segueing into a discussion of the directions of future work.

5.2 CONTRIBUTIONS AND INSIGHTS

Chapter 2 presented a novel framing of *non-root intra-paradigmatic alternations* from a diachronic perspective. Specifically, I provided a typology for these alternations based on their diachronic source, exemplified by several examples for each type. The main distinction is between alternations which were conditioned by phonological change and those which were created through morphological change. Of particular interest are IRREGULAR DISTRIBUTIONS. These alternations, also known as morphomic distributions, are characterized by a distribution which does not align with a natural class. As a result, these distributions pose difficulties from both diachronic and synchronic perspectives, as outlined by Chapter 2. The subsequent chapters investigated two examples of IRREGULAR DISTRIBUTIONS, examining how they interact with analogical change.

Chapter 3 examined a set of irregular verbs in Latin, which demonstrated several examples of non-root intra-paradigmatic alternations. The first part of the chapter investigated the synthetic-analytic distribution affecting the verbs *nōlō* ‘I want’ and *mālō* ‘I prefer’. I provided a systematic examination of the distribution of these forms in Plautus, resulting in a novel explanation for their distribution on the basis of metrical position. The first half of the chapter culminated in a diachronic account for the development of the verb *nōlō*, with a comparison between two potential pathways. The first, a phonological approach, assumes that periphrastic variants were simply inherited

from cases in which contraction did not take place due to a phonological restriction. The second is a morphological approach in which an uncontracted form in the 2SG is extended to the 3SG and 2PL on the basis of an IRREGULAR DISTRIBUTION affecting several other irregular verbs. While both accounts are plausible, I have argued in favor of a combination of the two, in which a phonological dispreference for contraction into heavy syllables allowed an already existing pattern to extend, namely the athematic-thematic distinction found in the present tense of other irregular verbs.

The second part of the chapter focused on the distribution of athematicity from a synchronic perspective. The motivation for accounting for the distribution using a synchronic framework is that presumably the extension of patterns to novel lexical items occurs during the acquisition process and thus is synchronic in nature (cf. Albright 2002a, Yang 2016, *a.o.*). Specifically, I compared two frameworks, DISTRIBUTION MORPHOLOGY (Halle & Marantz, 1993) and PARADIGM FUNCTION MORPHOLOGY (Stump, 2001, 2015, 2016). While PFM failed to account for the complicated distribution due to its emphasis on stems, the piece-based nature of DM allowed for a generalizable characterization of the athematic pattern. Building on Calabrese (2015b,a), Embick (2016), this account proposed that an impoverishment rule deleted the feature responsible for the insertion of the thematic vowel in the environment of the IRREGULAR DISTRIBUTION. In effect, the rule can simply be extended to novel items, motivating analogical change. Secondly, this comparison supported the fact that morphemes, (i.e. IRREGULAR DISTRIBUTIONS), can be accounted for using non-stem based analyses, contra a regularly assumed position in the literature (cf. Maiden 2018).

Finally, Chapter 4 investigated the interaction between a well-established morphomic distribution, the N-PATTERN, and analogical leveling. The N-PATTERN is a well-established distribution in which the present tense forms of the singular and 3PL differ

from those in the 1PL and 2PL. This pattern is found in many varieties of Romance, with several types of morphological phenomena obeying the distribution (cf. Maiden 2018); however, in some cases, the pattern has been eroded in various ways in favor of one of the alternants. I examined how this leveling occurred in several varieties, including languages in northern Italy and the region close to the Mediterranean spanning the area between Italy and Spain.

Specifically, I focused on the reflexes of *-sc-*, a suffix whose original function in Latin was inchoative but has since been generalized as a present tense marker. Ultimately, there are two pathways of development involving leveling of the alternants in the 1PL and 2PL on the basis of person and mood. I have argued that the developments that affect the instances of leveling were ultimately due to two factors, frequency (as conditioned by usage) and a desire for morphological alternations to express natural classes. As I show, changes involving person can be motivated by frequency in the sense that learners are less likely to uphold unique patterns if they have insufficient input. The changes in mood, namely the removal of the alternation from the present subjunctive can be explained via a paradigm-internal leveling which extends the marker of the present tense to all forms of the paradigm. Each of the attested individual changes demonstrate these characteristics, suggesting that learners are biased to change their input on the basis of both frequency and morphological cohesiveness.

I have also demonstrated that previous explanations for the individual instances of leveling are unable to account for the changes in a unified manner. Likewise, the traditional approach to the N-PATTERN, namely that it constitutes a template of sorts (Maiden, 2003, 2018), does not explain why the distributions level in certain directions. Taken together with the fact that this morpheme can be explained via impoverishment (cf. Embick 2016), it is unlikely that the N-PATTERN and its variants are best explained

via a templatic approach. The diachronic evidence further indicates that a templatic approach is ultimately untenable. Crucially, the verb ‘to go’ which regularly demonstrates a suppletive alternation in the shape of the N-PATTERN often lags behind other verbs in terms of leveling. Undergo a templatic approach, we would expect all verbs to behave alike, including the verb ‘to go’. This can easily be accounted for by frequency, as the verb ‘to go’ is simply more frequent in the 1PL and 2PL than other verbs, allowing for the maintenance of N-PATTERN alternants in these cells. In all, it is clear that morphomic distributions are susceptible to analogical change, putting into question their special status from a synchronic point of view.

5.3 DIRECTIONS FOR FURTHER WORK

The initial stated goal of this dissertation was to provide a basis for constraining analogy. While the investigation of IRREGULAR DISTRIBUTIONS and leveling biases do provide key insights into the limits of analogy, this should only be seen as the beginning. Of course, a full undertaking required to explain what constraints analogy is outside the scope of this dissertation. In this sense, there are several aspects discussed throughout this investigation which offer promising directions for further work.

There are several general ways in which the investigation can be broadened. From an empirical perspective, it is worth considering whether the same trends and biases hold if we examine both other language families with different morphological characteristics, as well as other types of morphological phenomena. Do other types of morphological expression have similar diachronic origins? Certainly suppletion exhibits many of the same characteristics (cf. [Kim 2019](#), [Juge 2019](#)). Additionally, more data regarding irregular distributions might prove insightful in terms of both their diachronic origin as well as their synchronic status.

The next few sections describe some more specific areas which deserve more attention.

5.3.1 MORPHOMES AND PRODUCTIVITY

This dissertation has examined several aspects of the intersection between analogy and morphemes, with special attention given to alternations involving non-root morphology. In doing so, I have provided several motivations for both why an irregular pattern might extend as well as how a pattern might change over time. For example, we can illustrate the first with the Latin example, where the morphomic pattern may have extended to *nōlō* in part due to a phonological restriction involving contraction into heavy syllables. As for the second, Chapter 4 examined how frequency and morphological biases directly impact which cells become leveled.

One aspect of analogy that has been mentioned in this dissertation is its interaction with productivity. While I have discussed some of the implications of this perspective, intra-paradigmatic alternations, specifically those with IRREGULAR DISTRIBUTIONS (i.e. morphomic), are particularly relevant. Consider, for example, the N-PATTERN. While the origin of this pattern clearly lies in a phonological alternation involving stress, its extension to other types of verbs is a bit of a mystery. From a descriptive perspective, the pattern was introduced to many different verb types, encompassing several types of morphological expression (suppletion, affixation, vowel alternations, etc.). Why was the pattern extended to some verb classes and not others? Productivity may provide an answer to these questions.

Consider the following scenario. The original distribution of the stress pattern would have affected certain conjugation classes, including the first conjugation which was (and often still is) the most productive verb class in Romance. If learners posited a

productive rule, such as an impoverishment rule affecting the 1PL and 2PL (cf. Embick 2016) using an algorithm like the Tolerance Principle (Yang, 2016), this rule could be applied to novel lexical items such as the verb ‘to go’. This impoverishment rule would align the expression in the 1/2PL with morphology that expresses the elsewhere form, accounting for suppletive and affixal alternations without requiring any additional assumptions. Thus, for Italian *andàre*, the creation of the *vad-/and-* alternation can be seen as a productive process in which the form in the 1PL and 2PL is brought in line with the other tenses (i.e. non-present). While this is a hypothetical, there is little reason why productivity could not be incorporated in our diachronic explanations of historical phenomena in a similar way in order to better explain why changes occurred the way they did.

The Latin pattern poses an interesting dilemma for productivity. Unlike the N-PATTERN, the athematic pattern is not found with a large selection of verbs; in fact only four verbs (not counting *nōlō* and *mālō*) show the thematic-athematic distinction. This might suggest that productivity does not play a part given how few verbs participate in the alternation. However, for the Tolerance Principle, it is possible that the rule responsible for the athematicity is still productive as long as the number of exceptions remains low (cf. Section 1.4.1.2). What is difficult for a productivity approach is determining which verbs the impoverishment rule applies to. Besides the fact that they engage in the athematic-thematic distinction, there is no obvious factor that unites *ferō*, *volō*, *sum*, and *edō*. It is unclear what type of verb the pattern could be extended to. In this sense, there is a question of how to formalize morphomic rules in such a way that a model can assess their productivity.

5.3.2 THEORY AND EXTENSION

A second issue which has not been fully dealt with in this dissertation involves several theory-specific points.

The first point pertains to the status of disjunctivity. Most notably in the Latin case, I have assumed that the impoverishment rule involves several disjunct environments. While doing so necessarily weakens the restrictiveness of the theory, I see no obvious solution to this problem given the distribution of athematicity. However, future research may investigate whether the environment can be stated in a unified way.

It should be noted that allowing disjunctivity is a topic that deserves discussion. Depending on the framework, different levels of disjunctivity make different predictions. Consider the distribution of athematicity in Latin. If we assume that we can have a singular rule listed with several disjunct environments, we might predict that any time the rule is extended to a novel item, it must do so in all of the disjunct environments. This would be equivalent to a templatic approach to morphomicity. If, however, we assume that the disjunct environments imply several identical rules that only differ in terms of which cells they apply to, the possibility arises that only a subset of the rules extend to a novel lexical item. This is essentially what the extension to *nōlō* implies. Must we simply accept that disjunctivity is a viable part of the grammar? If so, is there a way to constrain disjunctivity to correctly account for the data that we have, while not over-generating?

The second point concerns the extension of the athematic pattern to *nōlō*, whose distribution is characterized by periphrasis, not athematicity. The extension must be framed in terms of the environment, not the deletion of thematicity; in this sense, it is not the rule which is extended, but the environment. While Appendix A.6 discusses this issue in more depth, it is clear that there are some broader issues at stake. The Latin

data suggests that learners may consider extending the environment and not the specific rule. This implies that learners may have an incentive to reinforce the environments they encounter. A parallel can be found in Russian, where many of the nouns juxtapose the morphological expression in the plural with that in the singular (cf. Section 2.3.2.3). In this sense, there is an emergent pressure for paradigms to conform to a certain pattern. It remains to be seen whether this is a real phenomenon that affects the acquisitional process and thus diachronic change, or whether there is an alternative explanation for cases like the Latin pattern.



Latin Addendum

A.1 PARADIGMS

This section provides full paradigms for the verbs *volō* ‘I want’, *nōlō* ‘I do not want’, and *mālō* ‘I prefer’ in Classical Latin.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>volō</i>	<i>vīs</i>	<i>vult</i>	<i>volumus</i>	<i>vultis</i>	<i>volunt</i>
IMPF IND	<i>volēbam</i>	<i>volēbās</i>	<i>volēbat</i>	<i>volēbāmus</i>	<i>volēbātis</i>	<i>volēbant</i>
FUT	<i>volam</i>	<i>volēs</i>	<i>volet</i>	<i>volēmus</i>	<i>volētis</i>	<i>volent</i>
PERF IND	<i>voluī</i>	<i>voluistī</i>	<i>voluit</i>	<i>voluimus</i>	<i>voluistis</i>	<i>voluērunt</i>
PLFP IND	<i>volueram</i>	<i>voluerās</i>	<i>voluerat</i>	<i>voluerāmus</i>	<i>voluerātis</i>	<i>voluerant</i>
FUT PERF	<i>voluerō</i>	<i>volueris</i>	<i>voluerit</i>	<i>voluerimus</i>	<i>volueritis</i>	<i>voluerint</i>
PRES SUBJ	<i>velim</i>	<i>velīs</i>	<i>velit</i>	<i>velīmus</i>	<i>velītis</i>	<i>velint</i>
IMPF SUBJ	<i>vellem</i>	<i>vellēs</i>	<i>vellet</i>	<i>vellēmus</i>	<i>vellētis</i>	<i>vellent</i>
PERF SUBJ	<i>voluerim</i>	<i>voluerīs</i>	<i>voluerit</i>	<i>voluerīmus</i>	<i>voluerītis</i>	<i>voluerint</i>
PLPF SUBJ	<i>voluissem</i>	<i>voluissēs</i>	<i>voluisset</i>	<i>voluissēmus</i>	<i>voluissētis</i>	<i>voluissent</i>

Table A.1: Finite forms of *volō* ‘I want’ in Latin

(1) Infinite forms of *volō* ‘I want’ in Latin

a.	<i>volle</i>	PRES INF
b.	<i>voluisse</i>	PERF INF
c.	<i>volēns</i>	PRES PTPL

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>nōlō</i>	<i>nōn vīs</i>	<i>nōn vult</i>	<i>nōlumus</i>	<i>nōn vultis</i>	<i>nōlunt</i>
IMPF IND	<i>nōlēbam</i>	<i>nōlēbās</i>	<i>nōlēbat</i>	<i>nōlēbāmus</i>	<i>nōlēbātis</i>	<i>nōlēbant</i>
FUT	<i>nōlam</i>	<i>nōlēs</i>	<i>nōlet</i>	<i>nōlēmus</i>	<i>nōlētis</i>	<i>nōlent</i>
PERF IND	<i>nōluī</i>	<i>nōluistī</i>	<i>nōluit</i>	<i>nōluimus</i>	<i>nōluistis</i>	<i>nōluērunt</i>
PLPF IND	<i>nōlueram</i>	<i>nōluerās</i>	<i>nōluerat</i>	<i>nōluerāmus</i>	<i>nōluerātis</i>	<i>nōluerant</i>
FUT PERF	<i>nōluerō</i>	<i>nōlueris</i>	<i>nōluerit</i>	<i>nōluerimus</i>	<i>nōlueritis</i>	<i>nōluerint</i>
PRES SUBJ	<i>nōlim</i>	<i>nōlīs</i>	<i>nōlit</i>	<i>nōlīmus</i>	<i>nōlītis</i>	<i>nōlīnt</i>
IMPF SUBJ	<i>nōllem</i>	<i>nōllēs</i>	<i>nōllet</i>	<i>nōllēmus</i>	<i>nōllētis</i>	<i>nōllent</i>
PERF SUBJ	<i>nōluerim</i>	<i>nōluerīs</i>	<i>nōluerit</i>	<i>nōluerīmus</i>	<i>nōluerītis</i>	<i>nōluerint</i>
PLPF SUBJ	<i>nōluissem</i>	<i>nōluissēs</i>	<i>nōluisset</i>	<i>nōluissēmus</i>	<i>nōluissētis</i>	<i>nōluissent</i>
PRES IMP	–	<i>nōlī</i>	–	–	<i>nōlīte</i>	–
FUT IMP	–	<i>nōlītō</i>	<i>nōlītō</i>	–	<i>nōlītōte</i>	<i>nōluntō</i>

Table A.2: Finite forms of *nōlō* ‘I do not want’ in Latin

(2) Infinite forms of *nōlō* ‘I do not want’ in Latin

a.	<i>nōlle</i>	PRES INF
b.	<i>nōluisse</i>	PERF INF
c.	<i>nōlēns</i>	PRES PTPL

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>mālō</i>	<i>māvīs</i>	<i>māvult</i>	<i>mālumus</i>	<i>māvultis</i>	<i>mālunt</i>
IMPF IND	<i>mālēbam</i>	<i>mālēbās</i>	<i>mālēbat</i>	<i>mālēbāmus</i>	<i>mālēbātis</i>	<i>mālēbant</i>
FUT	<i>mālam</i>	<i>mālēs</i>	<i>mālet</i>	<i>mālēmus</i>	<i>mālētis</i>	<i>mālent</i>
PERF IND	<i>māluī</i>	<i>māluistī</i>	<i>māluit</i>	<i>māluimus</i>	<i>māluistis</i>	<i>māluērunt</i>
PLPF IND	<i>mālueram</i>	<i>māluerās</i>	<i>māluerat</i>	<i>māluerāmus</i>	<i>māluerātis</i>	<i>māluerant</i>
FUT PERF	<i>māluerō</i>	<i>mālueris</i>	<i>māluerit</i>	<i>māluerimus</i>	<i>mālueritis</i>	<i>māluerint</i>
PRES SUBJ	<i>mālim</i>	<i>mālīs</i>	<i>mālīt</i>	<i>mālīmus</i>	<i>mālītis</i>	<i>mālīnt</i>
IMPF SUBJ	<i>māllem</i>	<i>māllēs</i>	<i>māllet</i>	<i>māllēmus</i>	<i>māllētis</i>	<i>māllent</i>
PERF SUBJ	<i>māluerim</i>	<i>māluerīs</i>	<i>māluerit</i>	<i>māluerīmus</i>	<i>māluerītis</i>	<i>māluerint</i>
PLPF SUBJ	<i>māluissem</i>	<i>māluisēs</i>	<i>māluisset</i>	<i>māluissemus</i>	<i>māluissetis</i>	<i>māluisent</i>

Table A.3: Finite forms of *mālō* ‘I prefer’ in Latin

(3) Infinite forms of *mālō* ‘I prefer’ in Latin

a.	<i>mālle</i>	PRES INF
b.	<i>māluisse</i>	PERF INF
c.	<i>mālēns</i>	PRES PTPL

A.2 COUNTS OF *NŌLŌ* AND *MĀLŌ* IN PLAUTUS

A.2.1 *NŌLŌ*

A.2.1.1 PRESENT TENSE

	SINGULAR	#	PLURAL	#
1ST	<i>nōlō</i>	81	<i>nōlumus</i>	1
2ND	<i>nōn vīs</i>	9	<i>nōn voltis</i>	1
3RD	<i>nōn volt</i>	5	<i>nōlunt</i>	2

Table A.4: Regular forms of PRES IND

(4) Textual positions of notable forms

- a. ***nōn vīs* (x10):** *Amph.* 835, *Aul.* 161, *Cap.* 309, *Men.* 787-788, *Mil.* 682, *Most.* 336, *Pers.* 487, *Trin.* 1160, *Vid.* 29, [*Stich.* 483]
***nōn volt* (x5):** *Asin.* 460, *Cas.* 700, *Epid.* 70, *Epid.* 585, *Pers.* 613
***nōn voltis* (x1):** *Amph.* 56
- b. ***nevīs* (x10):** *Curc.* 82, *Merc.* 150, *Most.* 762, *Most.* 1176, *Pers.* 358, *Poen.* 1079, *Pseud.* 436, *Trin.* 328, *Trin.* 1156, *Truc.* 546
***nevolt* (x5):** *Epid.* 42, *Epid.* 586, *Most.* 110, *Trin.* 361, *Trin.* 364

FORM	#
<i>nōn volō</i>	2
<i>neque volō</i>	2
<i>neque vīs</i>	1
<i>neque volumus</i>	1
<i>nec volō</i>	1

Table A.5: Other negated forms in the PRES IND

(5) PRES SUBJ of *nōlō*

nōlim (x5): *Amph.* 86, *Amph.* 439, *Asin.* 214, *Cap.* 943, *Merc.* 539

nōlīs (x3): *Bacch.* 914, *Cas.* 314, *Pers.* 489

nōlit (x1): [*Most.* 287]

nōn velīs (x2): *Poen.* 244, *Trin.* 671

nōn velīt (x2): *Merc.* 452, *Merc.* 453

nōn velīnt (x2): *Merc.* 7, *Most.* 681

nec velīt (x1): *Curc.* 218

A.2.1.2 OTHER TENSES

	SINGULAR	#	PLURAL	#
1ST	<i>nōlēbam</i>	2	<i>nōlēbāmus</i>	0
2ND	<i>nōlēbās</i>	0	<i>nōlēbātis</i>	0
3RD	<i>nōlēbat</i>	0	<i>nōlēbant</i>	0

Table A.6: IMPF IND of *nōlō*

FORM	#
<i>nōllem</i>	x1 (<i>Stich.</i> 513)
<i>nōn vellem</i>	x1 (<i>Cist.</i> 506)
<i>nōn velles</i>	x1 (<i>Aul.</i> 286)

Table A.7: IMPF SUBJ of *nōlō*

	SINGULAR	#	PLURAL	#
1ST	<i>nōlam</i>	0	<i>nōlēmus</i>	0
2ND	<i>nōlēs</i>	2	<i>nōlētis</i>	0
3RD	<i>nōlet</i>	4	<i>nōlent</i>	0

Table A.8: FUT of *nōlō*

	SINGULAR	#	PLURAL	#
1ST	<i>nōlam</i>	3	<i>nōlēm̄us</i>	0
2ND	<i>nōlēs</i>	0	<i>nōlēt̄is</i>	0
3RD	<i>nōlet</i>	8	<i>nōlent</i>	0

Table A.9: PERF IND of *nōlō*

FORM	#
<i>nōn voluī</i>	2
<i>neque voluī</i>	1
<i>neque voluit</i>	2
<i>nec voluit</i>	1
<i>nōluerit</i>	1

Table A.10: Other negated perfects of *nōlō*

FORM	#
<i>nōluisset</i>	1
<i>nōlle</i>	3

Table A.11: Other negated forms of *nōlō*

A.2.2 MĀLŌ

A.2.2.1 PRESENT TENSE

	SINGULAR	#	PLURAL	#
1ST	<i>mālō</i>	7	<i>mālum̄us</i>	0
2ND	<i>māvīs</i>	7	<i>māvoltis</i>	1
3RD	<i>māvolt</i>	0	<i>mālunt</i>	0

Table A.12: Regular forms of PRES IND

(6) Textual positions of notable forms

***māvīs* (x7):** *Cas.* 290, *Mil* 216, *Most.* 177, *Poen.* 706, *Rud.* 853, *Stich.* 700, *Truc.* 153

***māvoltis* (x1):** *Bacch.* 1119

māvolō (x8): *Asin.* 835, *Curc.* 320, *Merc.* 421, *Poen.* 301, *Pseud.* 728, *Rud.* 1413, *Frag. Add.*1.1, [*Poen.* 303]

magis volō (x1): *Cist.* 299 **magis vīs (x1):** *Mil.* 1337

(7) PRES SUBJ of *mālō*

mālīm (x15): *Asin.* 811, *Asin.* 895, *Bacch.* 490, *Bacch.* 514, *Bacch.* 875, *Cap.* 858, *Epid.* 119, *Poen.* 289, *Poen.* 1184, *Poen.* 1214, *Trin.* 762, *Truc.* 260, *Vid.* 102, [*Bacch.* 465, *Bacch.* 519c]

mālīs (x2): *Amph.* 511, *Cist.* 33

mālīt (x1): [*Most.* 287]

mālint (x1): *Stich.* 80

māvelīm (x16): *Asin.* 877, *Aul.* 661, *Bacch.* 859, *Cap.* 516, *Men.* 720, *Merc.* 356, *Merc.* 889, *Pers.* 4, *Poen.* 151, *Poen.* 827, *Poen.* 1150, *Rud.* 570, *Truc.* 277, *Truc.* 422, *Truc.* 742, *Truc.* 746

māvelīs (x2): *Cap.* 270, *Pseud.* 140

māvelīt (x1): *Trin.* 306

A.2.2.2 OTHER TENSES

(8) Other Forms of *mālō*

a. IMPF SUBJ

māvellem (x9): *Amph.* 512, *Bacch.* 198, *Bacch.* 452, *Bacch.* 1047, *Bacch.* 1201, *Curc.* 512, *Mil.* 171, *Pseud.* 131, *Pseud.* 1057

b. FUTURE

māvolet (x1): *Asin.* 121

c. PERFECT

māluī (x1): *Mil.* 1356

A.3 SECONDARY FACTORS: USAGE AND FREQUENCY

Before discussing the synchronic analysis of the morphomic distribution in Latin, I briefly investigate whether usage may have played a role in determining which variants were used for the different cells of the paradigm. That is, it is worth considering whether the choice between a SYNTHETIC variant and SEMI-ANALYTIC/ANALYTIC variant was impacted by the usage of the two variants, either due to the frequency of the specific cells in question or due to the nature of the variation itself. There is a long line of research connecting patterns of contraction with usage (Bybee & Scheibman, 1999, Bybee, 2006b, Bresnan, 2021, *a.o.*) that may inform our understanding of this specific alternation. While ultimately, I do not argue in favor of a usage-based approach as the primary causal factor behind alternations of SEMI-ANALYTIC/ANALYTIC behavior in these verbs, it is nonetheless instructive to examine how such an explanation might reinforce the other mechanisms at play.

Following Bybee (2006a,b), there are two ways in which frequency may have played a role in affecting the choice of SYNTHETIC and SEMI-ANALYTIC/ANALYTIC variants. On the one hand, there is the REDUCING EFFECT, in which frequent forms are more likely to be reduced than less frequent forms, e.g. contraction in English auxiliaries *want to* → *wanna*, *do not* → *don't*. On the other hand, the CONSERVING EFFECT may cause more frequent forms to preserve morphological irregularities, e.g. less common irregular past tense verbs in English are more likely to regularize. While these effects are to my knowledge not normally applied intra-paradigmatically, we may make two hypotheses in which these effects may apply to *nōlō* and *mālō*, specifically in terms contraction and non-contraction.

The REDUCTION hypothesis for the Latin case can be stated as follows. If the forms of the 1SG *nōlō* < **nouolō*, 1PL *nolumus* < **nouolomus*, and 3PL *nolunt* < **nouolunt*

(those that resulted in SYNTHETIC forms) were more frequent than the other forms, they would be more likely to undergo contraction. These contracted forms would have then contrasted with the less frequent cells in the paradigm, which, due to being less frequent, would have remained non-univerbated in most instances. The alternate CONSERVATION hypothesis assumes that the forms of the 2SG $n\bar{o}n\ v\bar{i}s < nev\bar{i}s < *nou\bar{e}is$, 3SG $n\bar{o}n\ volt < nevolt < *nouolt$, and 2PL $n\bar{o}n\ voltis < *nevoltis < *nouoltis$ were more frequent than the other forms and were thus more likely to undergo and uphold the irregular ANALYTIC and SEMI-ANALYTIC patterns.¹

It is important to note that both hypotheses assume that there is some distinction between the frequencies of the two groups ($\{1SG, 1PL, 3PL\}$ versus $\{2SG, 2PL, 3SG\}$). For the REDUCTION hypothesis, it is the first group that is more frequent and therefore is contracted, while for the CONSERVATION hypothesis the second group is more frequent and therefore conserves morphological irregularity. Crucially, under either hypothesis, we expect that one of the two sets should be distinct from the other in terms of frequency. Due to this expected difference in frequency, if we simply compare the overall frequencies, we should be able to broadly determine whether frequency is a key causal factor of the distribution seen in the present indicative.

As we can see in the counts below for both *volō* (Table A.13), *nōlō* (Table A.14), and *mālō* (Table A.15) in Plautus,² no immediate pattern emerges that separates the two desired sets. That is, the two groups are not diametrically opposed to one another in terms of frequency for the three verbs in question.

¹The forms of *mālō* can be dealt with similarly in terms of both hypothesis *mutatis mutandis*.

²The counts in Table A.14 and Table A.15 not only include the regular SYNTHETIC vs SEMI-ANALYTIC/ANALYTIC alternations, but also cases in which a form of *volō* is tmetetic from either negation of the adjective *magis*.

	SINGULAR	PLURAL
1ST	x336	x11
2ND	x273	x22
3RD	x92	x25

Table A.13: PRES IND of *volō* in Plautus (Counts)

	SINGULAR	PLURAL
1ST	x86	x2
2ND	x19	x1
3RD	x10	x1

Table A.14: PRES IND *nōlō* in Plautus (Counts)

	SINGULAR	PLURAL
1ST	x15	—
2ND	x8	x1
3RD	—	—

Table A.15: PRES IND of *mālō* in Plautus (Counts)

The respective frequency of the specific cells do show a pattern, namely that the first singular followed by the second singular are clearly the most frequent of the forms. The plural forms are generally less frequent than the singulars for each of the verbs.

It is possible that the genre of text may influence these frequencies, and therefore the counts do not reflect the true synchronic distribution accessible to learners. In terms of written texts, the high frequency of the third person may be attributed to the genre, where third person is more common in prose and less common than the first and second person (specifically the singular) in the works of Plautus. Given that the texts may

bias the frequency, and one cannot directly test the acquisition of these forms in Latin, I compare the frequencies in Italian for the verb *volere* ‘to want’ for both child-directed and child speech from the Italian CHILDES corpus.³

The justification for using this corpus is twofold. Firstly, Italian is one of the languages that will best reflect the frequency of similar sequences in Latin, as it is a direct descendent. While the languages are distinct and are separated by two thousand years of diachronic change, I assume that the contexts in which utterances occur in the acquisition phase are not significant enough to prevent comparison. While it is certainly possible that we may use any language in principle to examine these frequencies, Italian is no less representative than any other. Secondly, Italian still conjugates its verbs for each of the persons and does not have syncretism unlike other languages, e.g. English. This allows us to compare the frequencies of the individual person and number combinations.⁴

Table A.16 and Table A.17 show the counts for the attested forms produced by speakers speaking to children and forms produced by children learning respectively. It is clear from this data that there is no pattern that singles out sets outlined above. Taken together with the frequencies of the respective verbs in Latin, it does not seem as if the grouping in Latin of the first person forms and the third plural versus the second person forms and the third singular can be due to the frequency of the individual cells.

Admittedly, it is still possible that the frequency may play a role; particularly notable is the high frequency of the first and second singular in both the Latin and Italian data.

³The data used here was acquired from the **Italian portion** of the CHILDES corpus (Antelmi, 1997, Morlacchi & Antelmi, 2005, Cipriani et al., 1989, D’Odorico et al., 2001, D’Odorico & Carubbi, 2003, Salerni et al., 2007, Antinucci & Parisi, 1973, Antinucci & Volterra, 1975, Volterra, 1976, Tonelli et al., 1995, 1998, Tonelli & Fabris, 2005).

⁴Regarding data collection, there are two relevant points. Firstly, Italian has no proper descendent of *nōlō* and otherwise has no contracted form for ‘do not want’. Some of the forms of *volere* ‘to want’ in the corpus are negated and are not counted separately. Secondly, I count the data from the investigators as well as any family members who were not the child under investigation as ‘child-directed’.

	SINGULAR		PLURAL	
1ST	<i>voglio</i>	x67	<i>vogliamo</i>	x30
2ND	<i>vuoi</i>	x1249	<i>volete</i>	x3
3RD	<i>vuole</i>	x183	<i>vogliono</i>	x27

Table A.16: ‘Child-directed’ forms of *volere* ‘to want’ in Italian

	SINGULAR		PLURAL	
1ST	<i>voglio</i>	x190	<i>vogliamo</i>	x3
2ND	<i>vuoi</i>	x44	<i>volete</i>	x1
3RD	<i>vuole</i>	x45	<i>vogliono</i>	x4

Table A.17: Child-uttered forms of *volere* ‘to want’ in Italian

If these cells served as ‘analogical anchors’ by which the other forms were modeled on the part of the learner, in this sense, frequency would play a role. For example, if a learner is mostly exposed to the forms in the singular (as these are the most frequent), they could use this information to fill in the paradigm, particularly when combining this data with the abstract generalization that the 2SG, 2PL, and 3SG cohere. In all, however, frequency does not seem to be the major causal factor in deriving the relevant pattern.⁵

⁵As discussed above, the variance between SYNTHETIC and SEMI-ANALYTIC/ANALYTIC forms was largely determined by emphasis. In this sense, it is worth considering the intersection between emphasis and usage. If there is a reason to distinguish between the two variants in terms of emphasis, specifically between the two sets of cells in question, then we can attribute the distribution of the two sets to the difference in emphatic usage. In this sense, we would expect either the set {2SG, 2PL, 3SG} or {1SG, 1PL, 3PL} to be used in more emphatic contexts than the other.

Ultimately, however, such a pattern is not apparent from the texts. For example, the distribution of the emphatic particle *ēdepol* ‘by god’ co-occurring with the negative particle *nōn* is largely limited to the first person singular, followed by the second person singular. There are also examples of the first plural, albeit significantly less common than the first and second singular:

- (i) *nōn ēdepol faciēmus*
NEG by.Pollux do.FUT.1PL
‘by god, we will not do [it]!’ (Pl. *Epid.* 263)

It is not immediately clear how emphasis would be limited to certain person and number combinations. Even though it is relatively easy to relate negation to emphasis, e.g. *-n’t* versus *not* in English, that no obvious usage can be identified suggests that emphasis did not play a specific role in this respect.

A.4 VOCABULARY ITEMS

A.4.1 VI RULES FOR *FER* $\bar{O}$:

(9) VOCABULARY ITEMS for *fer* $\bar{o}$ ‘I bear’

- a. $v_{\text{bear}} \leftrightarrow \text{tul} / \text{ ______ } [+perf]$
- b. $v_{\text{bear}} \leftrightarrow \text{fer}$

A.4.2 VI RULES FOR *VOL* $\bar{O}$:

(10) VOCABULARY ITEMS for *vol* $\bar{o}$ ‘I want’

- a. $v_{\text{want}} \leftrightarrow \text{vel} / \text{ ______ } [-perf, +subj], [+pres, +infinite]$
- b. $v_{\text{want}} \leftrightarrow \text{v}\bar{\text{i}} / \text{ ______ } [+2, +sg, +pres, +ind]$
- c. $v_{\text{want}} \leftrightarrow \text{volu} / \text{ ______ } [+perf]$
- d. $v_{\text{want}} \leftrightarrow \text{vol}$

A.4.3 VI RULES FOR *SUM*:

The last two verbs, *sum* ‘I am’ and *ed* $\bar{o}$ ‘I eat’, are of special interest as the patterned athematicity is accompanied directly by root suppletion. The environments of the vocabulary items must therefore reflect this fact in combination with the impoverishment rule that deletes the [+TV] diacritic as laid out in (22) and (23). The impoverishment rule allows the vocabulary items responsible for the ‘morphomic’ root alternates *es* and *$\bar{e}s$* for *sum* and *ed* $\bar{o}$ respectively to be underspecified. Specifically, they must be underspecified in terms of the [+TV] diacritic, allowing for insertion of *es* and *$\bar{e}s$* as the elsewhere case.

Beginning with *sum* ‘I am’, the VIs for this verb are listed below in (11).

(11) VOCABULARY ITEMS for *sum* ‘I am’

- a. $v_{be} \leftrightarrow s / ___ [+pres, +TV]$
- b. $v_{be} \leftrightarrow fu / ___ [+perf]$
- c. $v_{be} \leftrightarrow es$

As with *ferō*, the distinction between the infectum and the perfectum holds for this verb, allowing the allomorph *fu* to be handled straightforwardly. A more difficult part of this set of rules is the specification for the environment for the allomorph *s*. While the present indicative forms that show *s-* are unproblematic, the present subjunctive forms, as seen in Table A.18, present an issue.

	CLASSICAL	PLAUTUS
1SG	<i>sim</i>	<i>siem</i>
2SG	<i>sīs</i>	<i>siēs</i>
3SG	<i>sit</i>	<i>siet</i>
1PL	<i>sīmus</i>	<i>sīmus</i>
2PL	<i>sītis</i>	<i>sītis</i>
3PL	<i>sint</i>	<i>sient</i>

Table A.18: PRES SUBJ for *sum* ‘I am’ in Latin

Table A.18 shows the paradigm for the present subjunctive as well as some of the variants seen in Old Latin. The Plautine variants are notably archaic and show the old ablaut pattern of the optative marker PIE $*_{\check{\iota}}eh_1-$ / $*_{\check{\iota}}ih_1-$. However, as pointed out by Weiss (2020: 453), the ablauting variants are limited to verse-final position.⁶

The difficulty that these forms pose is that the lack of theme vowel in these forms indicates that they should be spelled out by the elsewhere case, *es*, if the spell-out rules are formulated as in (11). If the present subjunctive of *esse* is expounded by *s-*, the root should be marked by the [+TV] diacritic, which in turn should insert a theme vowel.

⁶See the distribution of verse-final and medial variants for the SEMI-ANALYTIC, ANALYTIC, and SYNTHETIC forms discussed in Section 3.2.3.

There are several possible solutions, but I suggest that the problem is independent of the specific exponence of *esse*. As we can see in Table A.19, the present subjunctive for the regular conjugations is not uniform with regard to the exponence of the theme vowel.

	1ST (- <i>ā</i> -)	2ND (- <i>ē</i> -)	3RD (- <i>e</i> -)	3RD - <i>iō</i> (- <i>i</i> -)	4TH (- <i>i</i> -)
	ĀMŌ ‘love’	VIDEŌ ‘see’	DŪCŌ ‘lead’	CAPĪŌ ‘take’	AUDIŌ ‘hear’
1SG	<i>āmēm</i>	<i>videam</i>	<i>dūcam</i>	<i>capiam</i>	<i>audiam</i>
2SG	<i>āmēs</i>	<i>videās</i>	<i>dūcās</i>	<i>capiās</i>	<i>audiās</i>
3SG	<i>āmet</i>	<i>videat</i>	<i>dūcat</i>	<i>capiat</i>	<i>audiat</i>
1PL	<i>āmēmus</i>	<i>videāmus</i>	<i>ducāmus</i>	<i>capiāmus</i>	<i>audiāmus</i>
2PL	<i>āmētis</i>	<i>videātis</i>	<i>dūcātis</i>	<i>capiātis</i>	<i>audiātis</i>
3pl	<i>āment</i>	<i>videant</i>	<i>dūcant</i>	<i>capiant</i>	<i>audiant</i>

Table A.19: PRES SUBJ forms in Latin

In the first and third conjugations, the theme vowel is not present, suggesting that the lack of a theme vowel in the PRES SUBJ of *sum* (and also *edō*) is not an isolated irregularity. Whatever is responsible for the lack of insertion of the theme vowel in the PRES. SUBJ. of the first and third conjugations may ultimately also be responsible for the lack of theme vowel in the irregular verbs. I leave the exact answer to why the theme vowel is missing in these forms aside, as it is clearly an issue relevant to the exponence of the present subjunctive as a whole, rather than the specific vocabulary items for *sum*.⁷

Returning to the VI rules for *sum*, if we ignore the issue related to the present subjunctive, the attested forms in the present indicative that show *es* are derived directly from the combination of the VI rules and the impoverishment rule.

⁷A potential solution that may account for the lack of theme vowel in the third conjugation and the irregular verbs is phonological. As the theme vowel in the third conjugation (cf. *fer-i-mus*) and the irregular verbs (e.g. *s-u-mus/vol-u-mus*) is short, it would be deleted before the long vowel representing the present subjunctive, *ā* and *ī* for the third conjugation and irregular verbs respectively. This deletion would result in the attested forms. It does not explain the lack of theme vowel in the first conjugation as that vowel is underlyingly long: *ā*. It should be noted that the PRES. SUBJ. 1SG in the first conjugation has the same discrepancy: *āmō* ‘I love’ (instead of **āmaō*).

An added benefit to positing the form *es* for the elsewhere case is that it is also the form of the root that appears in the imperfect and future. This allows us to account for the generalization that the exponent *es* is not exclusively a marker for athematicity; instead, the elsewhere form is used in many forms of the verb, a subset of which are those patterned for athematicity.

VI RULES FOR *EDŌ*: The VIs for the final irregular verb, *edō* ‘I eat’, largely mirror those of *sum*, as seen in (12).

(12) VOCABULARY ITEMS for *edō* ‘I eat’

- a. $v_{\text{eat}} \leftrightarrow \text{ed} / \text{ ______ } [-\text{perf}, [+TV]]$
- b. $v_{\text{eat}} \leftrightarrow \bar{\text{ed}} / \text{ ______ } [+perf]$
- c. $v_{\text{eat}} \leftrightarrow \bar{\text{es}}$

The distinction between the forms in the perfectum and infectum is generally maintained in this verb by the alternation of *ed-* (infectum) and *ēd-* (perfectum). The form *ed* refers to the [+TV] diacritic in the same way as the present allomorph for *esse* does. It should be noted that the future and imperfect tenses of the third conjugation do not show a theme vowel across the board; as with the present subjunctive, I assume this is the result of some independent phenomenon. After the relevant forms have been impoverished in terms of the theme vowel, we are left with the elsewhere form *ēs*.

A.5 STEM FUNCTIONS

A.5.1 STEM FUNCTIONS FOR THE VERB *EDŌ*:

The verb *EDŌ* ‘I eat’ exhibits a direct correlation between athematicity and root suppletion. Here I use the capitalized form to represent the abstract lexeme that stands

for this verb. The relevant **Stem** functions in Latin that pertain to this root are given below in (13).

(13) Clauses in the **Stem** function in Latin relevant to EDŌ ‘I eat’

- a. **Stem**($\langle \bar{E}SSE, \sigma : \{\text{imperfective}\} \rangle$) = *ede*
- b. **Stem**($\langle \bar{E}SSE, \sigma : \{\text{perfective}\} \rangle$) = *ēd*
- c. **Stem**($\langle \bar{E}SSE, \sigma : \{\text{pret imperfective}\} \rangle$) = *edēb*
- d. **Stem**($\langle \bar{E}SSE, \sigma : [\{\text{imp subj}\} \vee [\{\text{pres}\} \wedge [\{\text{inf}\} \vee [\{2\} \vee \{3 \text{ sg}\}]]]] \rangle$) = *ēs*
- e. **Stem**($\langle \bar{E}SSE, \sigma : \{\text{pres 3 pl}\} \rangle$) = *edu*

The stems of this verb are straightforward. Conforming with other Latin verbs, the imperfective and perfective stems correspond generally to the forms in the perfectum and the infectum.⁸ I have also posited a stem for the imperfect tense, following the distribution of the ‘*b*-stem’ set up in [Stump \(2015\)](#). Assuming the non-decomposition of the theme vowel, the stem that produces the form 3 PL PRES IND *edunt* must be handled as a separate stem as it does not follow directly from any synchronic phonological process. Finally, the stem that is associated with athematicity is encoded with a disjunct environment, following the system of set-theoretic formulations used in the framework (cf. [Stump 2015: 46](#)).⁹

A.5.2 STEM FUNCTIONS FOR THE VERB SUM:

The **Stem** functions that pertain to SUM ‘I am’ are listed below in (14):

⁸For the proposed stem functions, I have assumed that the theme vowel can be deleted in the 1SG PRES IND *edō* before the long vowel. A similar approach can be taken for the future.

⁹Arguably, the **Stem** function in (13) is quite charitable towards PFM. Above, I have accounted for the forms in the imperfect subjunctive with the stem *ēs*. However, according to the STEM MAXIMISATION PRINCIPLE, the stem for the imperfect subjunctive should be *ēssē*, not *ēs*: *ēssem*, *ēssēs*, *ēsset*, *ēssēmus*, *ēssētis*, *ēssent*. In this sense, the imperfect subjunctive is not athematic inherently, as it is just another listed stem.

(14) Clauses in the **Stem** function in Latin relevant to SUM ‘I am’

- a. **Stem**($\langle \text{ESSE}, \sigma : \{\text{imperfective}\} \rangle$) = *es*
- b. **Stem**($\langle \text{ESSE}, \sigma : \{\text{perfective}\} \rangle$) = *fu*
- c. **Stem**($\langle \text{ESSE}, \sigma : \{\text{pres subj}\} \rangle$) = *sī*
- d. **Stem**($\langle \text{ESSE}, \sigma : [\{\text{pres ind}\}] \wedge [\{1\} \vee \{3 \text{ pl}\}] \rangle$) = *su*

As with EDŌ, I have posited a distinction between imperfective and perfective stems, expounded for SUM by *es* and *fu* respectively.¹⁰ The stems for the present subjunctive and the PRES IND 1SG/1PL/3PL are split in this analysis, contra the DM-style approach. This bifurcation is justified by the fact that otherwise the vocalic segment present in the PRES IND, *-u-* (e.g. *s-u-mus*, *s-u-nt*) would have to be analyzed as an independent thematic vowel, unlike the other thematic vowels for the regular classes.

It should be noted at this point that the environments which specify the stems that realize the athematic cells do not overlap for the two verbs discussed so far. For ĒSSE, there is a specific stem that targets the root allomorph that aligns with the athematic pattern, namely *ēs*. For SUM, however, the stem that is realized as *es* must cover a larger set of cells (beyond those that inflect athematically), including the cells of the imperfect and future tenses.¹¹ In other words, nothing in the framework so far has made specific reference to athematicity or the morphomic pattern.

¹⁰This analysis differs from the DM analysis in that the stem *es* is specified for imperfective rather than as an elsewhere case stem. This distinction is ultimately immaterial as the environment *imperfective* is functionally equivalent to an elsewhere case, given that the forms in the present subjunctive and present indicative that do not have *es* as their stem are specified in their respective stem functions.

¹¹The imperfect subjunctive falls under the purview of the stem *es* in the **Stem** function in (14) (cf. fn. 9). This could, and perhaps should, be analyzed with the stem *essē*. This further demonstrates the lack of cohesiveness of a stem-based analysis of athematicity.

A.6 EXTENSION TO $n\bar{o}l\bar{o}$

The distribution of periphrasis in $n\bar{o}l\bar{o}$ poses an interesting question for the synchronic accounts discussed in Section 3.4. Specifically, the DM account assumes that an impoverishment rule accounts for the presence of athematicity in the paradigms of *volō*, *ferō*, *edō*, and *sum*. However, the alternation in $n\bar{o}l\bar{o}$ (and $m\bar{a}l\bar{o}$) is not really expressed in terms of athematicity. It is true that the periphrastic forms are athematic; however, this seems to be more of a side effect of being a derivative of *volō*. In other words, it is the presence of *volō* in the derivation that provides the athematicity in these cells. In this sense, how does the impoverishment rule stated below explain the extension to $n\bar{o}l\bar{o}$?

- (15) Delete [+TV] in the context $\{[+pres, +ind, -part, -pl]_{AGR}, [+pres, +ind, +part, -auth]_{AGR}, [+impf, +subj]_T, [+pres, +inf]_T\}$

I suggest that it is not deletion of the theme vowel expressed in (15) which is extended to $n\bar{o}l\bar{o}$. Rather, it is the environment (or a subset thereof) which is extended to another impoverishment rule.

In order to express this notion, consider how we might account for the periphrasis in $n\bar{o}l\bar{o}$. While we might approach this problem from several angles, I propose that contraction is governed by an abstract diacritic [+M], which expresses whether merger occurs.¹² The cells which have this feature contract, while those that lack the feature do not contract. In order to account for the distribution of contraction, we can invoke the following impoverishment rule:

¹²This diacritic is inspired by a similar feature proposed by Bobaljik (2012).

- (16) Delete [+M] in the context $\{[+pres, +ind, -part, -pl]_{AGR}, [+pres, +ind, +part, -auth]_{AGR}\}$

The vocabulary insertion rules can thus reference the presence (and absence) of the diacritic. The forms which have the [+M] diacritic have *nōl-*, while those that do not allow for the spell out of the regular marker for negation (i.e. *nōn*).

This account assumes that a learner specifically learns that the 2SG, 2PL, and 3SG are periphrastic. The alternative, namely that a learner memorizes that the 1SG, 1PL, and 3PL are contracted, is a logical possibility which might seem preferable given the irregularity of contraction. Most verbs in Latin are not contracted when negated; the forms in the 2SG, 2PL, and 3SG are the expected outcomes of the combination of *nōn* with *volō*. However, I argue that this alternative obscures the general distribution of contraction in the paradigm. As can be seen in Appendix A.1, it is only the 2SG, 2PL, and 3SG of the PRES IND which are not contracted. We would have to assume that every cell apart from these was explicitly targeted by a feature insertion rule, a solution which does not seem particularly economical. Additionally, the Plautine forms, *nevī s* and *nevolt*, do not suggest a simple combination of the negation with the verb *volō*, as the general marker of negation was *nōn*. In this sense, it is clear that the cells in the 2SG, 2PL, and 3SG are targeted by the impoverishment rule.

B

Romance Addendum

This section provides background on the Romance verbal system, beginning with how it differs from its predecessor.

B.1 BACKGROUND ON LATIN VERBAL SYSTEM

This section summarizes the relevant synchronic facts surrounding the categorical organization of the verbal system as well as the different verbal conjugations.

B.1.1.1 CATEGORICAL ORGANIZATION

The Latin verbal system contains many different categories that reflect tense, aspect, and mood. While there has been much discussion about how exactly these categories should be described, I lay out the traditional explanation here. Going back to the Roman grammarian Varro, the verbal morphology in Latin is organized around two grammatical aspects (Weiss 2020: 421). The *infectum* aspect refers to ongoing actions, irrespective of tense. The *perfectum*, on the other hand, refers to completed actions. The different tenses and moods can be subdivided into these two categories. B.1 demonstrates this division in terms of tense.

	PRES	PST	FUTURE
INFECTUM	present	imperfect	future
PERFECTUM	perfect	pluperfect	future perfect

Table B.1: *infectum* and *perfectum* split in Latin

This division is largely motivated by the fact that the verbal morphology often reflects the difference. For example, consider the suppletion in the verb *ferō* ‘I bear’, as seen in the first person singular indicative forms in B.2:

	PRES	PST	FUT
INFECTUM	<i>ferō</i> ‘I am bearing’	<i>ferēbam</i> ‘I was bearing’	<i>feram</i> ‘I will bear’
PERFECTUM	<i>tulī</i> ‘I have brought’	<i>tuleram</i> ‘I had brought’	<i>tulerō</i> ‘I will have brought’

Table B.2: Suppletive split for *ferō* ‘I bear’ in Latin

In these forms, the form *fer-* alternates with *tul-* along the lines of the distinction between the *infectum* and *perfectum*.

In summary, Latin separates its verbal system into two grammatical aspects.

B.1.1.1 DISTRIBUTION OF -SC-

These aspects have morphological consequence also affected the distribution of the suffix -sc-. As discussed in Section 4.2.1, the suffix -sc- was limited to the *infectum*. Thus, the suffix was found in the present, imperfect, and future, while it is absent from the perfect. It is not restricted by mood and occurs in the indicative, subjunctive, and imperative. Consider the paradigm for *calēscere* ‘to become hot’:

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>calēscō</i>	<i>calēscis</i>	<i>calēscit</i>	<i>calēscimus</i>	<i>calēscitis</i>	<i>calēscunt</i>
IMPF IND	<i>calēscēbam</i>	<i>calēscēbās</i>	<i>calēscēbat</i>	<i>calēscēbāmus</i>	<i>calēscēbātis</i>	<i>calēscēbant</i>
FUT	<i>calēscam</i>	<i>calēscēs</i>	<i>calēscet</i>	<i>calēscēmus</i>	<i>calēscētis</i>	<i>calēscant</i>
PRES SUBJ	<i>calēscam</i>	<i>calēscās</i>	<i>calēscat</i>	<i>calēscāmus</i>	<i>calēscātis</i>	<i>calēscant</i>
PRES IMP	—	<i>calēsce</i>	—	—	<i>calēscite</i>	—
IMPF SUBJ	<i>calēscerem</i>	<i>calēscerēs</i>	<i>calēsceret</i>	<i>calēscerēmus</i>	<i>calēscerētis</i>	<i>calēscerent</i>

Table B.3: *calēscō* ‘I become hot’ in Latin

B.1.2 LATIN CONJUGATION

As with many other Indo-European languages, Latin has different categories into which verbs fall when it comes to their morphological expression. These categories are expressed through four main conjugations with one sub-conjugation. These conjugations are characterized by a *theme vowel*, a vocalic segment which differs in terms of vowel quality and length depending on which class the verb expresses. The present indicative of the different conjugation classes are shown below in B.4, alongside the relevant theme vowel for each class.

The four main theme vowels, alongside the -*iō* type, are expressed between the root and the personal endings, although there are some differences which I do not discuss in detail here.

Under many synchronic analyses (e.g. [Aronoff 1994](#), [Embick 2000](#), [Stump 2015](#),

	1ST (-ā-)	2ND (-ē-)	3RD (-e-)	3RD -iō (-i-)	4TH (-ī-)
	ĀMĀRE ‘love’	VIDĒRE ‘see’	DŪCERE ‘lead’	CAPERRE ‘take’	AUDĪRE ‘hear’
1SG	<i>āmō</i>	<i>videō</i>	<i>dūcō</i>	<i>capiō</i>	<i>audiō</i>
2SG	<i>āmās</i>	<i>vidēs</i>	<i>dūcis</i>	<i>capis</i>	<i>audīs</i>
3SG	<i>āmāt</i>	<i>videt</i>	<i>dūcit</i>	<i>capit</i>	<i>auditt</i>
1PL	<i>āmāmus</i>	<i>vidēmus</i>	<i>ducimus</i>	<i>capimus</i>	<i>audīmus</i>
2PL	<i>āmātis</i>	<i>vidētis</i>	<i>dūcitis</i>	<i>capitis</i>	<i>audītis</i>
3PL	<i>āmant</i>	<i>vident</i>	<i>dūcunt</i>	<i>capiunt</i>	<i>audiunt</i>

Table B.4: Overview of conjugation in Latin

a.o.), the theme vowel is assumed to have no inherent value and merely serves to signal inflectional class membership. Some analyses therefore do not decompose the theme vowel at all and treat it simply as part of the stem, while others consider it as a suffix that simply expresses morphological information, namely class membership. However, membership in the different conjugation classes is not always arbitrary, and many significant trends can be identified. For example, the different conjugations can largely be associated with certain grammatical functions, some of which are listed below.¹

- 1ST CONJUGATION: denominal verbs, either from first declension nouns, or from adjectives, e.g. *cūrāre* ‘to care’ (alongside *cūra* ‘care’), *novāre* ‘to renew’ (alongside *novus* ‘new’).
- 2ND CONJUGATION: Stative verbs, e.g. *calēre* ‘to be hot’, *tacēre* ‘to be quiet’
- 4TH CONJUGATION: denominal verbs from *i*-stems: *finīre* ‘to finish’ (alongside *finis* ‘end’)

Alongside these example can be found several suffixes associated with certain declension classes, such as the repetitive *-it-* (first conjugation), *-ig-* (first conjugation), *-ic* (first conjugation), and *-sc-* (third conjugation).

¹These examples were retrieved from Weiss (2020).

To summarize the conjugation system, there are four (and a half) conjugations into which Latin verbs can be classified. Class membership is not necessarily arbitrary and suffixes are associated with certain conjugations.

B.2 ROMANCE VERBAL SYSTEM

B.2.1 RESTRUCTURING OF LATIN MORPHOLOGICAL CATEGORIES

This section briefly provides background of the general restructuring of the verbal system on the way from Latin to the modern languages. I do not discuss every single formation here and disregard passives entirely, for purposes of brevity.²

The general trend in the development of the Romance verb was to eliminate synthetic forms with analytic forms (cf. Meul 2016: 41, Weiss 2020: 563). In the *inflectum*, the synthetic future was replaced by an analytical future which in most languages was a combination of the infinitive with the present tense of *habēō* ‘I have’, e.g. It. *canterò* ‘I will sing’ < *cantāre habēō* replacing Lat. *cantābō* ‘I will sing’. A similar construction involving the combination of the infinitive with the *imperfect* tense of *habēō* has given rise to the conditional. In Italian, the conditional was formed by the perfect tense of *habēō* instead of the imperfect: It. *canterebbe* ‘(s)he would sing’ < Lat. *cantāre habuīt*.

The rest of the *inflectum* is preserved. Thus, the present indicative and subjunctive are synthetic in the modern languages, as is the imperfect indicative.

In the *perfectum*, the perfect with a past function was largely preserved as the preterite, with the different Latin perfect types undergoing different fates (cf. Weiss 2020: 567-572). The other indicative tenses in the *perfectum*, the pluperfect and future perfect are more restricted and have various outcomes in Western Romance, with the latter restricted to Ibero-Romance and Gascon (Weiss 2020: 573). The perfect subjunc-

²For more detailed discussions, see Alkire & Rosen (2010), Meul (2016, *a.o.*).

tive (as well as the imperfect subjunctive) are largely irrelevant for future discussion, as their outcomes are heavily restricted in the modern languages.

The most relevant subjunctive (besides the present subjunctive) is the outcome of the pluperfect subjunctive, which in Western Romance has become the imperfect subjunctive, e.g. It. *cantasse* < Lat. *cantāvisset*. The *-ss-* that characterizes the imperfect subjunctive much resembles the outcome of *-sc-* in the modern languages (cf. Section 4.6.1). However, it should be noted that these suffixes are distinct in terms of function and accordingly have different distributions. The imperfect subjunctive suffix is occasionally found in the conditional, as is the case in the Lombard dialect Paves *finärisä* ‘(s)he would finish’.

I summarize the aforementioned changes in the table below, ignoring some of the tenses and moods from the *perfectum* that have specific difficulties and are not relevant to discussion in upcoming sections:

INFECTUM	PERFECTUM
PRES IND >PRES IND	PERF IND >PRET IND
PRES SUBJ >PRES SUBJ	
IMPF IND >IMPF IND	PLPF SUBJ >IMPF SUBJ
INF + <i>habeō</i> >FUT	
INF + <i>habēbam/habuī</i> > COND	

Table B.5: Relevant categorical changes from Latin to Modern Romance

As can be seen in B.5, in some ways the difference between the *infectum* and *perfectum* are still maintained; however, the introduction of new constructions like the conditional as well as the reanalysis of the pluperfect subjunctive as the imperfect subjunctive complicates the original division between the two categories.

B.2.2 ROMANCE CONJUGATION CLASS

The different modern Romance languages largely preserve a system of verbal classification in which class membership is signaled by the presence of a theme vowel. For illustration, consider the verbal system as described in [Oltra-Massuet \(2000\)](#) for Catalan. There are three conjugation classes with the third divided into two subclasses, seen below in the gerund forms in (1).³

- (1) Catalan conjugation classes
- a. Class I: *cant-a-nt* ‘singing’ (-a-)
 - b. Class II: *tem-e-nt* ‘fearing’ (-e-)
 - c. Class IIIa: *un-i-nt* ‘joining’ (-i-)
 - Class IIIb: *surt-i-nt* ‘going out’ (-i-)

However, the Romance and Latin conjugations differ in two ways. The first difference is that in Romance, exemplified here by Catalan, membership to the different classes is almost entirely arbitrary. The first conjugation is the most productive of these classes, and newly formed verbs, regardless of function, generally belong to this class. This is not to say that there are no associations between function and class. For example, some members of the third class are associated with the deadjectival notion of ‘cause to become X’ ([Oltra-Massuet, 2000](#)), similar to the inchoative use of -*sc-* in Latin. Besides this small subgroup, however, there are no obvious trends present in the conjugation classes connecting the theme vowels with a syntactico-semantic function.

This situation is common with the other Romance languages and has contributed to the notion that theme vowels are largely empty segments (cf. [Aronoff \(1994\)](#), [Oltra-](#)

³These examples were retrieved from [Oltra-Massuet \(2000\)](#). The gerund conveniently shows the difference in theme vowel between the different classes.

Massuet (2000), Maiden (2018), *a.o.*). In this sense, the Romance conjugation system can be seen as a ‘semantically bleached’ version of the Latin system. While the general distribution of the conjugations is still present, much of the functionality of the different theme vowels has been eroded, leaving what are verb classes that are separate from any obvious morphosyntactic conditioning.

The second way in which the Romance verbal system differs in terms of the conjugation classes is that the four (and a half) declension types seen in Latin have largely been reduced to three declension classes in many of the daughter languages. To begin, in all Romance languages, the third *-iō* conjugation has been lost as a unique class and merged with the fourth conjugation (cf. Meul 2016: 39).

More significant changes affected the second and third conjugations. The differences between these two conjugations have eroded in many Romance languages, although not in a unified manner. In most languages, the second and third conjugation only differ in a few forms like the infinitive. For example, Italian *rèndere* ‘to give’ shows stress on the root as expected for a third conjugation verb (cf. Lat. *reddĕre* ‘to give back’), but the stress on a second conjugation verb is not on the root, e.g. Italian *vedĕre* ‘to see’ < Lat. *vidĕre* ‘to see’ (cf. Meul 2016: 40). This is accompanied in many languages by movement from one conjugation to another by individual lexical items.

Ibero-Romance is unique in that it completely merges the second and third conjugations, resulting in three conjugations that largely cohere to the first, second, and fourth conjugations in Latin (Meyer-Lübke, 1894, Alkire & Rosen, 2010, *a.o.*). Some languages, like Romanian have preserved all four conjugation systems, although these four classes are obscured by the many additional subclasses that have developed, including those involving the descendants of verbal suffixes *-sc-* and *-idi-* (cf. Maiden 2017).

Because of these developments, individual Romance languages often have what are

essentially three conjugations, rather than four. As a result, the standard descriptive labels for these conjugations may differ depending on how many there are in a given language.

B.2.3 ASSUMED SUBGROUPING

Below, I lay out the genetic relationships that I assume, following [Weiss \(2020\)](#) among others. The status of certain controversial groupings like Catalan with either Ibero-Romance or Gallo-Romance are not of significant importance for our purposes, and nothing in this dissertation hinges upon which grouping one uses.

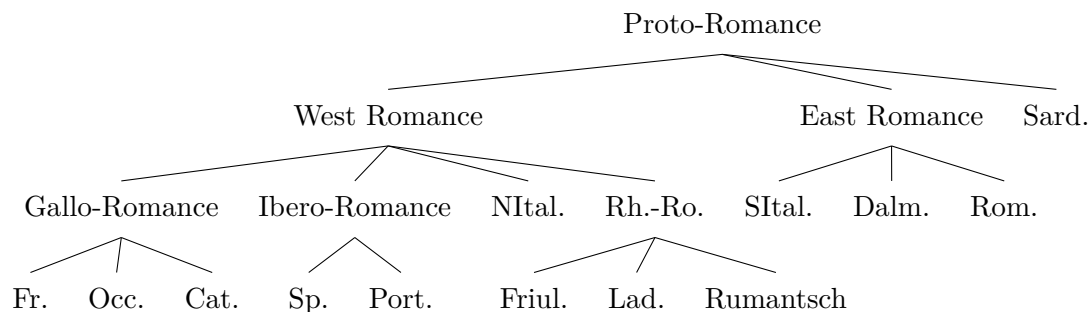


Figure B.1: Romance languages, following [Weiss \(2020\)](#)

B.3 NON-PATTERNED DISTRIBUTIONS

B.3.0.1 IBERO-ROMANCE LANGUAGES

The first NON-PATTERNED distribution under discussion is found in Ibero-Romance. The distribution in Spanish and Portuguese presents no intra-paradigmatic variation of the suffix, and is found throughout entire present indicative and subjunctive as well as the rest of the paradigm, including the forms descended from the *perfectum*. For example, consider the subset of the paradigms for Spanish *padeecer* ‘to suffer’ and Portuguese

padecer ‘to suffer’ in Table B.6 and Table B.7 respectively.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>padezco</i>	<i>padeces</i>	<i>padece</i>	<i>padecemos</i>	<i>padecéis</i>	<i>padecen</i>
IMPF IND	<i>padecía</i>	<i>padecías</i>	<i>padecía</i>	<i>padecíamos</i>	<i>padecíais</i>	<i>padecían</i>
PRET IND	<i>pedecí</i>	<i>padeciste</i>	<i>padeció</i>	<i>padecimos</i>	<i>padecisteis</i>	<i>padecieron</i>
PRES SUBJ	<i>padezca</i>	<i>padezcas</i>	<i>padezca</i>	<i>padezcamos</i>	<i>padezcáis</i>	<i>padezcan</i>

Table B.6: *padecer* ‘to suffer’ in Spanish

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>padeço</i>	<i>padeces</i>	<i>padece</i>	<i>padecemos</i>	<i>padeceis</i>	<i>padecem</i>
IMPF IND	<i>padecia</i>	<i>padecias</i>	<i>padecia</i>	<i>padecíamos</i>	<i>padecíeis</i>	<i>padeciam</i>
PRET IND	<i>pedeci</i>	<i>padeceste</i>	<i>padeceu</i>	<i>padecemos</i>	<i>padecestes</i>	<i>padeceram</i>
PRES SUBJ	<i>padeça</i>	<i>padeças</i>	<i>padeça</i>	<i>padeçamos</i>	<i>padeçais</i>	<i>padeçam</i>

Table B.7: *padecer* ‘to suffer’ in Portuguese

The forms in these paradigms show that the reflex of *-sc-* in Spanish and Portuguese (*-ec < -ēsc-*) are found throughout the entire paradigm of these verbs.⁴ This distribution suggests three key observations pertaining to the development of *-sc-*: i) the suffix has moved to the second conjugation rather than the fourth (*-ec-* not **-ic-*), ii) the suffix has been extended to the forms descended from the *perfectum* (e.g. the preterite forms Sp. *pedecí*, Port. *pedeci*), and most importantly iii) there is no alternation between the presence of the suffix and absence of the suffix in any of the tense and mood combinations, specifically the PRES IND or the PRES SUBJ.

The first two points are two ways that Ibero-Romance departs from other Romance languages. The suffix *-sc-* in Latin was limited to the *infectum*, so expectedly most Romance languages do not show forms with the suffix in the tenses/moods descended from the *perfectum*. The extension of the suffix to the preterite (and other forms like the imperfect subjunctive) is therefore an innovation of the Ibero-Romance languages. This

⁴It should be noted that there is a pronunciation difference in Spanish between *-zc-* and *-c-* based on the quality of the following vowel.

development, in my view, can be seen as the consequence of the collapse of the third conjugation into the second conjugation. Once *-sc-* verbs became second conjugation rather than third, the suffix was reanalyzed as a marker for the second conjugation and was extended to the rest of the paradigm, accompanying the breakdown of the apposition between the *infectum* and *perfectum*.

The third point begs the question of whether Spanish and Portuguese inherited a PATTERNED distribution and simply leveled out any alternation in favor of the forms with *-sc-* or whether these languages never had a PATTERNED alternation to begin with. This question has produced two opposing views in the literature, those that assume a distribution, specifically the N-PATTERN, and those that assume there was no distribution. The data from the early periods of Spanish and Portuguese are inconclusive (cf. Meul 2016: 101-102). There is limited evidence that the infix alternated in the present tense in the N-PATTERN (Dworkin, 1985), which has led some to suppose that the alternation was leveled. However, it has also been suggested that the forms with and without the suffix represent a competition between stems with *-ecer* and those with *-ir*. Meul (2016), following evidence from Kauffeld (2007), argues that the difference between the forms with *-sc-* and those without really was a difference between finite forms (expressed with *-ec*) and infinite forms. This alternation was eventually leveled out in favor of forms entirely with *-ec* or without. Under such an analysis, there likely never was an alternation of forms with *-ec* and those without in Ibero-Romance; instead these languages likely inherited the regular distribution as seen in Latin, where the suffix was limited to the *infectum*.

For the purposes of this discussion, I assume the latter approach. In this view, Spanish and Portuguese, as we will see with French, inherited a system in which *-sc-* was limited to the *infectum*.

B.3.1 FRENCH

The French paradigm for the verb *finir* ‘to finish’ shows that the suffix is found throughout the forms descended directly from the infectum, seen below in B.8.⁵ The segmented sequences show a segment *-i-/-iss-* intervening between the root and the desinence. It is this segment that is the modern descendent of Proto-Romance **-isc-* in French. The singular forms, despite having only *-i-* also continue the same suffix (cf. Section 2.3.1.2). In this sense, all forms below can be derived from *-iss-*.

	1SG	2SG	3SG	1PL	2PL	3PL
PRES IND	<i>fin-i-s</i>	<i>fin-i-s</i>	<i>fin-i-t</i>	<i>fin-iss-ons</i>	<i>fin-iss-ez</i>	<i>fin-iss-ent</i>
IMPF IND	<i>fin-iss-ais</i>	<i>fin-iss-ais</i>	<i>fin-iss-ait</i>	<i>fin-iss-ions</i>	<i>fin-iss-iez</i>	<i>fin-iss-aient</i>
PRES SUBJ	<i>fin-iss-e</i>	<i>fin-iss-es</i>	<i>fin-iss-e</i>	<i>fin-iss-ions</i>	<i>fin-iss-iez</i>	<i>fin-iss-ent</i>

Table B.8: Finite forms with *-iss* < **-isc-* in French

The other finite cells of verbs of this class lack any reflex of *-iss-*. Instead, the *-iss-* is seemingly replaced synchronically by the regular fourth conjugation marker *-i-*. The preterite and the imperfect subjunctive were derived from the perfect and the pluperfect subjunctive respectively; the lack of *-iss-* in these forms is expected given the fact that it never appeared in the *perfectum*.⁶

The forms in the future and the conditional are more difficult. While the modern French paradigm lacks the suffix *-iss-* in these forms, several forms in Old French point to an intrusion: *floristrat* ‘it will bloom’, *esclarcistra* ‘it will illuminate’, etc.⁷ This is also demonstrated in some dialects of Walloon: 1PL FUT IND *fin-ih-rans* ‘we will finish’ (compare with Mod. French *finirons* ‘we will finish’ without the suffix).⁸ The

⁵These forms follow from their Latin counterparts. See Section B.1 for more information.

⁶The *-iss-* in the imperfect subjunctive in French is not from *-sc-*, but from the pluperfect subjunctive marker *-iss-* in Latin, e.g. *fīnīvissem* ‘I would have had finished’.

⁷cf. Meul (2016: 108) for more examples.

⁸cf. Lempereur & Morayns (1974). This dialect also seems to have extended the stem with *-iss-* (*-ih-*) to the perfect, similarly to Parisian French. In this sense, it is not clear whether the future and

future and conditional themselves are uncontroversially the result of the infinitival stem appended to the present and imperfect of *habeō* ‘I have’ respectively. In this sense, the expected outcome of fourth conjugation verbs with *-isc-* is represented by the forms in Modern French: **finīre + habēmus > finirens* ‘we will have’. In effect, the fact that *-sc-* does not appear in these forms is completely expected, given the source of these forms. It should also be noted that some other Romance dialects similarly have an intrusion of the suffix in the future and conditional, albeit not in the imperfect indicative (Maiden 2003: 32, Meul 2016: 112).

The distribution in French has several interesting implications for the development of *-sc-* in Romance. While the suffix does not follow the N-PATTERN (or any other PATTERNED distribution for that matter) in the present indicative and subjunctive, the language does have evidence for the N-PATTERN elsewhere, seen below in B.9, in the present indicative for the verb *aller* ‘to go’ and in phonologically-contrastive stem allomorphs of various verbs.

	1SG	2SG	3SG	1PL	2PL	3PL
ALLER ‘go’	<i>vais</i>	<i>vas</i>	<i>va</i>	<i>allons</i>	<i>allez</i>	<i>vont</i>
MOURIR ‘die’	<i>meurs</i>	<i>meurs</i>	<i>meurt</i>	<i>mourons</i>	<i>mourez</i>	<i>meurent</i>
DEVOIR ‘owe’	<i>dois</i>	<i>dois</i>	<i>doit</i>	<i>devons</i>	<i>devez</i>	<i>doivent</i>

Table B.9: PRES IND of verbs with N-PATTERN in French

In terms of analogy, this development presents us with two potential options. The first option is that the forms with the suffix *-iss-* < **-isc-* never had the N-PATTERN in the first place. The second option is that the alternant in the singular and 3PL was extended to all forms while the pattern itself was preserved in the verbs in B.9.

I assume the former approach, following both Meul (2016: 102-103) and Maiden (2018: 180). This assumption is grounded in the fact that an analogical spread of

conditional forms should be seen as representative of a leveling of the stem throughout the paradigm or as a preservation of an earlier distribution

the suffix to precisely the cells that descended from the *infectum* is more costly than assuming that French is simply more conservative in its distribution.

Bibliography

- Agarotti, C. (2005). Il mantovano. In A. Rognoni (Ed.), *Grammatica dei dialetti della Lombardia*. Milan: Mondadori.
- Aguirre, C. (2003). Early verb development in one Spanish-speaking child. In D. Bittner, W. Dressler, & M. Kilani-Schoch (Eds.), *Development of verb inflection in first language acquisition: A cross-linguistic perspective* (pp. 1–25). Berlin: De Gruyter Mouton.
- Ainsworth, W. & Paliwal, K. (1984). Correlation between the production and perception of the English glides /w, r, l, j/. *Journal of Phonetics*, 12(3), 237–243.
- Alberti, R. (1993). *Die Mundart von Gavardo (Prov. Brescia)*. Geneva: Droz.
- Albright, A. (2002a). *The identification of bases in morphological paradigms*. PhD thesis, Los Angeles: University of California, Los Angeles.
- Albright, A. (2002b). Islands of reliability for regular morphology: Evidence from Italian. *Language*, 78(4), 684–709.

- Albright, A. (2005). The morphological basis of paradigm leveling. In L. J. Downing, T. Hall, & R. Raffelsiefen (Eds.), *Paradigms in phonology* (pp. 17–43). Oxford: Oxford University Press.
- Albright, A. (2009). Modeling analogy as probabilistic grammar. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar* (pp. 185–213). Oxford: Oxford University Press.
- Albright, A. (2010). Base-driven leveling in Yiddish verb paradigms. *Natural Language & Linguistic Theory*, 28(3), 475–537.
- Albright, A. & Hayes, B. (2002). Modeling English past tense intuitions with minimal generalization. In M. Maxwell (Ed.), *Proceedings of the ACL-02 workshop on morphological and phonological learning* (pp. 58–69).: Stroudsburg, PA: ACL Publications.
- Alkire, T. & Rosen, C. (2010). *Romance languages: A historical introduction*. Cambridge: Cambridge University Press.
- Andersen, H. (1980). Morphological change: Towards a typology. In J. Fisiak (Ed.), *Historical morphology* (pp. 1–50). New York: Mouton.
- Anderson, S. R. (2008). Phonologically conditioned allomorphy in the morphology of Surmiran (Rumantsch). *Word Structure*, 1(2), 109–134.
- Anglade, J. (1921). *Grammaire de l'ancien provençal: ou ancienne langue d'oc*. Paris: Klincksieck.
- Antelmi, D. (1997). *La prima grammatica dell'italiano: indagine longitudinale sull'acquisizione della morfosintassi italiana*. Bologna: Il Mulino.

- Antinucci, F. & Parisi, D. (1973). Early language acquisition: A model and some data. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 607–619). New York, NY: Holt, Rinehart & Winston.
- Antinucci, F. & Volterra, V. (1975). Lo sviluppo della negazione nel linguaggio infantile: uno studio pragmatico. *Lingua e stile*, 10(2), 231–260.
- Aronoff, M. (1994). *Morphology by itself*. Cambridge, MA: MIT Press.
- Aronoff, M. (2016). Unnatural kinds. In A. R. Luís & R. Bermúdez-Otero (Eds.), *The Morphome debate* (pp. 11–32). Oxford: Oxford University Press.
- Austin, J. (2013). Markedness, input frequency, and the acquisition of inflection: Evidence from Basque/Spanish bilingual children. *International Journal of Bilingualism*, 17(3), 259–283.
- Azaretti, E. (1977). *L'evoluzione dei dialetti liguri: esaminata attraverso la grammatica storica del ventimigliese*. Sanremo: Edizioni Casabianca.
- Baayen, H. (1989). *A corpus-based approach to morphological productivity: Statistical analysis and psycholinguistic interpretation*. PhD thesis, Amsterdam: Vrije Universiteit.
- Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 109–149). Dordrecht: Kluwer.
- Baayen, H. (1993). On frequency, transparency and productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1992* (pp. 181–208). Dordrecht: Kluwer.

- Baayen, H. (2009). Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics: An international handbook* (pp. 900–919). Berlin & New York: De Gruyter Mouton.
- Bauer, L. (2001). *Morphological productivity*. Cambridge: Cambridge University Press.
- Bauer, L. (2008). Lenition revisited. *Journal of Linguistics*, 44(3), 605–624.
- Benincà, P. (2000). Lombardy. In G. Price (Ed.), *Encyclopedia of the languages of Europe* (pp. 261–262). Oxford: Blackwell.
- Bennett, C. E. (1898). *Critique of some recent subjunctive theories*. Ithaca, NY: Andrus & Church.
- Birjulin, L. A. & Xrakovskij, V. S. (2001). Imperative sentences: Theoretical problems. In V. S. Xrakovskij (Ed.), *Typology of imperative constructions* (pp. 3–50). Munich: Lincom Europa.
- Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42(3), 531–573.
- Blevins, J. P. (2016). *Word and Paradigm Morphology*. Oxford: Oxford University Press.
- Blevins, J. P., Ackerman, F., & Malouf, R. (2019). Word and Paradigm Morphology. In J. Audring & F. Masini (Eds.), *The Oxford handbook of morphological theory* (pp. 265–284). Oxford: Oxford University Press.
- Bobaljik, J. D. (2002). Syncretism without paradigms: Remarks on Williams 1981, 1994. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2001* (pp. 53–85). Dordrecht: Foris.

- Bobaljik, J. D. (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.
- Bonami, O. & Stump, G. T. (2016). Paradigm Function Morphology. In A. Hippisley & G. Stump (Eds.), *The Cambridge handbook of morphology* (pp. 449–481). Cambridge: Cambridge University Press.
- Börjars, K. & Vincent, N. (2011). The pre-conditions for suppletion. In A. Galani, G. Hicks, & G. Tsoulas (Eds.), *Morphology and its Interfaces* (pp. 239–265). Amsterdam: John Benjamins.
- Bresnan, J. (2021). Formal grammar, usage probabilities, and auxiliary contraction. *Language*, 97(1), 108–150.
- Burzio, L. (2005). Sources of paradigm uniformity. In L. J. Downing, T. Hall, & R. Faffelsiefen (Eds.), *Paradigms in phonological theory* (pp. 65–106). Oxford: Oxford University Press.
- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, J. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology* (pp. 119–141). New York: Academic Press.
- Bybee, J. (1991). Natural morphology: The organization of paradigms and language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories*, volume 2 (pp. 67–92). Philadelphia: John Benjamins.
- Bybee, J. (2006a). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

- Bybee, J. (2006b). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733.
- Bybee, J. & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics*, 37(4), 575–596.
- Calabrese, A. (2012). Allomorphy in the Italian passato remoto: A Distributed Morphology analysis. *Language and Information Society*, 18, 1–75.
- Calabrese, A. (2015a). Irregular morphology and athematic verbs in Italo-Romance. *Isogloss: Special Issue on Italo-Romance Morphosyntax*, (pp. 69–102).
- Calabrese, A. (2015b). Locality effects in Italian verbal morphology. In E. D. Domenico, C. Hamann, & S. Matteini (Eds.), *Structures, strategies and beyond: Studies in honor of Adriana Belletti* (pp. 97–133). Amsterdam: John Benjamins.
- Campbell, L. (2013). *Historical linguistics: An Introduction*. Cambridge, MA: MIT Press.
- Castellani, F. E. (2002). *Grammatica del dialetto di Cascinagrossa*. Alessandria: Edizioni dell'Orso.
- Chitoran, I. (2002). The phonology and morphology of Romanian diphthongization. *Probus*, 14(2), 205–246.
- Cipriani, P., Pfanner, P., Chilosì, A., Cittadoni, L., Ciuti, A., Maccari, A., Pantano, N., Pfanner, L., Poli, P., Sarno, S., Bottari, P., & Cappelli, G. (1989). *Protocolli diagnostici e terapeutici nello sviluppo e nella patologia del linguaggio*. Pisa: Italian Ministry of Health, Stella Maris Foundation.

- Clark, E. (1985). The acquisition of Romance, with special reference to French. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 702–782). Hillsdale, NY: Erlbaum.
- Comoletti, C. (2005). Il milanese e i dialetti della Lombardia sudoccidentale (pavese e lodigiano). In A. Rognoni (Ed.), *Grammatica dei dialetti della Lombardia*. Milan: Mondadori.
- Corbett, G. G. (2007). Canonical typology, suppletion, and possible words. *Language*, 83(1), 8–42.
- Coveney, A. (2000). Vestiges of *nous* and the 1st person plural verb in informal spoken French. *Language Sciences*, 22(4), 447–481.
- Cowgill, W. (1978). The source of Latin *vīs* ‘thou wilt’. *Die Sprache*, 24, 25–44.
- Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow: Longman.
- Crowley, T. & Bower, C. (2010). *An introduction to historical linguistics*. Oxford: Oxford University Press.
- Dahl, Ö. (1979). Typology of sentence negation. *Linguistics*, 17(1-2), 79–106.
- De Belder, M. & Van Craenenbroeck, J. (2015). How to merge a root. *Linguistic Inquiry*, 46(4), 625–655.
- Delbrück, B. (1871). *Der Gebrauch des Conjunctivs und Optativs im Sanskrit und Griechischen*. Halle: Waisenhaus.
- Delbrück, B. (1893). *Einleitung in das Sprachstudium*. Leipzig: Breitkopf & Härtel, 3rd edition.

- DiFabio, E. G. (1990). *The morphology of the verbal infix /-isk-/ in Italian and in Romance*. PhD thesis, Cambridge, MA: Harvard University.
- D’Odorico, L. & Carubbi, S. (2003). Prosodic characteristics of early multi-word utterances in Italian children. *First Language*, 23(1), 97–116.
- D’Odorico, L., Carubbi, S., Salerni, N., & Calvo, V. (2001). Vocabulary development in Italian children: A longitudinal evaluation of quantitative and qualitative aspects. *Journal of Child Language*, 28(2), 351–372.
- Durham, C. L. (1901). *The subjunctive substantive clauses in Plautus: Not including indirect questions*. Ithaca, NY: Andrus & Church.
- Dworkin, S. N. (1985). From *-ir* to *-ecer* in Spanish: the loss of OSp. de-adjectival *-ir* verbs. *Hispanic Review*, 53(3), 295–305.
- Embick, D. (2000). Features, syntax, and categories in the Latin perfect. *Linguistic Inquiry*, 31(2), 185–230.
- Embick, D. (2010). *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Embick, D. (2016). Separation and its limits. In A. R. Luís & R. Bermúdez-Otero (Eds.), *The Morphome debate* (pp. 276–305). Oxford: Oxford University Press.
- Embick, D. (2021). The motivation for roots in Distributed Morphology. *Annual Review of Linguistics*, 7(1), 69–88.
- Embick, D. & Halle, M. (2005). On the status of stems in morphological theory. In T. Geerts, I. van Ginneke, & H. Jacobs (Eds.), *Romance languages and linguistic theory 2003* (pp. 37–62). Amsterdam: John Benjamins.

- Ezeizabarrena, M. J. (1996). *Adquisición de la morfología verbal en euskera y castellano por niños bilingües*. Bilbao, Spain: Servicio Editorial de la Universidad del País Vasco.
- Fábregas, A. (2017). Theme vowels are verbs. *Rivista di Grammatica Generativa*, 39, 79–89.
- Fenne, F.-Z. E. (2020). Paradigm structure in French verbal inflection. *Scripta*, 24(51), 103–135.
- Fertig, D. (1998). Suppletion, natural morphology, and diagrammaticity. *Linguistics*, 36(6), 1065–91.
- Fertig, D. (2013). *Analogy and morphological change*. Edinburgh: Edinburgh University Press.
- Flechia, G. (1876). *Intorno ad una peculiarità di flessione verbale in alcuni dialetti lombardi: nota letta alla Reale Accademia dei Lincei il 18 giugno 1876*. Rome: Salviucci.
- Fortson, B. (2008). *Language and rhythm in plautus*. Berlin: Walter de Gruyter.
- Fortson, B. (2011). *Indo-European language and culture: An introduction*. Oxford: John Wiley & Sons, 2nd edition.
- Garrett, A. (2008). Paradigmatic uniformity and markedness. In J. Good (Ed.), *Linguistic universals and language change* (pp. 125–143). Oxford: Oxford University Press.
- Gönczöl, R. (2020). *Romanian: An essential grammar*. New York: Routledge, Taylor & Francis Group.
- Goryczka, P. (2022). Allomorphy conditioned by post-linearization spanning: Evidence from Italian theme elements. In M. Janebová, J. Emonds, & L. Veselovská

(Eds.), *Language use and linguistic structure: Proceedings of the Olomouc linguistics colloquium 2021* (pp. 50–65). Olomouc: Palacký University.

Gouskova, M. & Bobaljik, J. D. (2022). The lexical core of a complex functional affix: Russian baby-diminutive *-onok*. *Natural Language and Linguistic Theory*, 40(4), 1075–1115.

Grafström, Å. (1969). On remplaçant *nous* en français. *Revue de linguistique romane*, 33(131-132), 270–298.

Gragera, A. F. (2000). *The role of typological markedness in the acquisition of Spanish subjunctive and in language change*. PhD thesis, Amherst, MA: University of Massachusetts Amherst.

Grestenberger, L. (2022). To *v* or not to *v*? Theme vowels, verbalizers, and the structure of the Ancient Greek verb. *Glossa: a journal of general linguistics*, 47(1).

Griva, G. (1980). *Grammatica della lingua piemontese*. Turin: Viglongo.

Grosso, M. (2002). *Grammatica della lingua piemontese: norme essenziali di grafia e morfologia*. Turin: Nòste Rèis.

Hale, W. G. (1895). *The anticipatory subjunctive in Greek and Latin*. Chicago: University Press of Chicago.

Halle, M. & Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 111–176). Cambridge, MA: MIT Press.

- Halle, M. & Marantz, A. (1994). Some key features of Distributed Morphology. In A. Carnie, H. Harley, & T. Bures (Eds.), *Papers on Phonology and morphology* (pp. 275–288). Cambridge MA: MIT Working Papers in Linguistics.
- Harðarson, J. A. (1993). *Studien zum urindogermanischen Wurzelarist: und dessen Vertetung im Indoiranischen und Griechischen*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Harley, H. (2008). When is a syncretism more than a syncretism? Impoverishment, metasyncretism, and underspecification. In D. Harbour, D. Adger, & S. Béjar (Eds.), *Phi theory: Phi-features across Modules and Interfaces* (pp. 251–294). Oxford: Oxford University Press.
- Harley, H. (2014). On the identity of roots. *Theoretical Linguistics*, 40(3-4), 225–276.
- Harley, H. & Noyer, R. (1999). Distributed Morphology. *Glott International*, 4(4), 3–9.
- Harley, H. & Ritter, E. (2002a). Person and number in pronouns: A feature-geometric analysis. *Language*, 78(3), 482–526.
- Harley, H. & Ritter, E. (2002b). Structuring the bundle: A universal morphosyntactic feature geometry. In H. J. Simon & H. Wiese (Eds.), *Pronouns: Features and representation* (pp. 23–39). Amsterdam: John Benjamins.
- Hayes, B. P. & Albright, A. (1999). An automated learner for phonology and morphology. Manuscript, Los Angeles, CA: UCLA.
- Herman, J. (2000). *Vulgar Latin*. Pennsylvania: Pennsylvania State University Press.
- Hewson, J. & Bubenik, V. (2006). *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam: John Benjamins.

- Hill, E. (2007). Proportionale Analogie, paradigmatischer Ausgleich und Formerweiterung: Ein Beitrag zur Typologie des morphologischen Wandels. *Diachronica*, 24(1), 81–118.
- Hock, H. H. (1991). *Principles of historical linguistics*. Berlin: De Gruyter Mouton.
- Hock, H. H. & Joseph, B. D. (2009). *Language history, language change, and language relationship*. Berlin: De Gruyter Mouton.
- Huber, J. (1958). Zur Verbalflexion der Mundart von Livigno. *Vox Romanica*, 17(1), 82–128.
- Hyams, N. M. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.
- Iliescu, M. & Mourin, L. (1991). *Typologie de la morphologie verbale romane, Vue synchronique*. Innsbruck: Amoe.
- Isac, D. (2015). *The morphosyntax of imperatives*. Oxford: Oxford University Press.
- Janda, R. D. (2017). “Phonologization” as the start of dephoneticization—or, on sound change and its aftermath: Of extension, generalization, lexicalization, and morphologization. In R. D. Janda & B. D. Joseph (Eds.), *The handbook of historical linguistics* (pp. 401–422). Oxford: Blackwell.
- Jasanoff, J. H. (2003). *Hittite and the Indo-European verb*. Oxford: Oxford University Press.
- Jasanoff, J. H. (2019). The sigmatic forms of the Hittite verb. *Indo-European Linguistics*, 7(1), 13–71.

- Jaskuła, K. (2014). Intervocalic elision of labials in Polish. In E. Cyran & J. Szpyra-Kozłowska (Eds.), *Crossing phonetics-phonology lines* (pp. 63–74). Newcastle upon Tyne: Cambridge Scholars Press.
- Jaworski, S. (2009). Inertial and non-inertial lenition processes. *Poznań Studies in Contemporary Linguistics*, 45(1), 103–129.
- Jaworski, S. (2021). Phonetic reduction of intervocalic [w] in Contemporary Polish. *Studies in Polish Linguistics*, 16(2), 79–97.
- Jaworski, S. & Gillian, E. (2011). On the phonetic instability of the Polish rhotic/r. *Poznań Studies in Contemporary Linguistics*, 47(2), 380.
- Jespersen, O. (1917). *Negation in English and other languages*. Copenhagen: A.F. Høst.
- Juge, M. L. (1999). On the rise of suppletion in verbal paradigms. In S. S. Chang, L. Liaw, & J. Ruppenhofer (Eds.), *Proceedings of the twenty-fifth Annual Meeting of the Berkeley Linguistics Society*, volume 25 (pp. 183–194).: Berkeley, CA: Berkely Linguistics Society.
- Juge, M. L. (2013). Analogy as a source of suppletion. In R. K. . L. Reid (Ed.), *Historical linguistics 2011: Selected papers from the 20th International Conference on Historical Linguistics, Osaka, 25–30 July, 2011*. (pp. 175–197). Amsterdam: John Benjamins.
- Juge, M. L. (2019). The sense that suppletion makes: towards a semantic typology on diachronic principles. *Transactions of the Philological Society*, 117(3), 390–414.
- Juilland, A. & Chang-Rodríguez, E. (1964). *Frequency dictionary of Spanish words*. The Hague, Netherlands: De Gruyter Mouton.

- Kauffeld, C. (2007). Old Spanish *-ir/-ecer* verb variation: Tracing the extension of *-ec-* through the lexicon. *Bulletin of Spanish Studies*, 84(8), 965–987.
- Kaufman, J. H. (2011). *The acquisition of the subjunctive mood by intermediate-level learners of Spanish: The relationship between mood and modality*. PhD thesis, Phoenix: Arizona State University.
- Kim, R. I. (2019). Suppletive stem distribution: A diachronic perspective. *Transactions of the Philological Society*, 117(3), 455–470.
- Kiparsky, P. (1968). *How abstract is phonology?* Bloomington: Indiana University Linguistic Club. Reprinted in O. Fujimura (ed.) (1973) *Three dimensions of linguistic theory*, 5–56, Tokyo: TEC and in Kiparsky (1982a), 119–63.
- Kiparsky, P. (1992). Analogy. In W. Bright (Ed.), *International encyclopedia of linguistics, vol. 1* (pp. 56–61). New York: Oxford University Press.
- Kodner, J. (2019). Estimating child linguistic experience from historical corpora. *Glossa: A Journal of General Linguistics*, 4(1).
- Kodner, J. (2020). *Language acquisition in the past*. PhD thesis, University of Pennsylvania.
- Kodner, J. (2022). What learning Latin verbal morphology tells us about morphological theory. *Natural Language & Linguistic Theory*, 41, 1–60.
- Kramer, R. (2016). Syncretism in paradigm function morphology and distributed morphology. In D. Siddiqi & H. Harley (Eds.), *Morphological metatheory* (pp. 95–120). Amsterdam: John Benjamins.

- Kuryłowicz, J. (1964). *The inflectional categories of Indo-European*. Heidelberg: Carl Winter.
- Kuryłowicz, J. (1966). La nature des procès dits «analogiques». In E. P. Hamp, F. W. Householder, & R. Austerlitz (Eds.), *Readings in linguistics II*, volume 5 (pp. 158–74). Chicago: Chicago University Press. [Reprinted from *Acta Linguistica* 5 (1945), 121–38].
- Kühlert, N. T. (2021). ‘Compensatory diphthongization’ in Ancient Greek dialects. Paper presented at the East Coast Indo-European Conference XL.
- Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world’s languages*. Oxford: Blackwell Publishers.
- Lampitelli, N. & Ulfsbjörninn, S. (2023). A classless analysis of Italian nouns and their theme-vowel alternations. *Isogloss. Open Journal of Romance Linguistics*, 9(1), 1–36.
- Langslow, D. (2009). *Jacob Wackernagel, Lectures on syntax: With special reference to Greek, Latin, and Germanic*. Oxford: Oxford University Press.
- Lanly, A. (1977). *Morphologie historique des verbes français*. Paris: Bordas.
- Lass, R. (1997). *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lausberg, H. (1962). *Romanische Sprachwissenschaft: Formenlehre: zweiter Teil*. Berlin: De Gruyter.
- Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Lempereur, J. & Morayns, J. (1974). *Li walon d’Lîdge. Tôme III: La conjugaison*. Liège: Imprimerie Wagelmans.

- Leumann, M., Hofmann, J. B., & Szantyr, A. (1977). *Lateinische Grammatik*, volume 2. CH Beck.
- Lieber, R. (1980). *On the organization of the lexicon*. PhD thesis, Cambridge, MA: Massachusetts Institute of Technology.
- Lisker, L. (1957). Minimal cues for separating /w, r, l, y/ in intervocalic position. *Word*, 13(2), 256–267.
- LIV² (2001). *Lexikon der indogermanischen Verben: Die Wurzeln und ihre Primärstammbildungen Unter der Leitung von Helmut Rix und der Mitarbeit vieler anderer bearbeitet von Martin Kümmel, Thomas Zehnder, Reiner Lipp, Brigitte Schirmer. Zweite, erweiterte und verbesserte Auflage bearbeitet von Martin Kümmel und Helmut Rix*. Wiesbaden: Reichert.
- Maddieson, I. & Emmorey, K. (1985). Relationship between semivowels and vowels: Cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica*, 42(4), 163–174.
- Maiden, M. (2003). Verb augments and meaninglessness in early Romance morphology. *Studi di grammatica italiana*, 22, 1–61.
- Maiden, M. (2004). When lexemes become allomorphs - On the genesis of suppletion. *Folia Linguistica*, 38(3-4), 227–256.
- Maiden, M. (2005). Morphological autonomy and diachrony. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 2004* (pp. 137–175). Dordrecht: Springer.
- Maiden, M. (2012). A paradox? The morphological history of the Romance present. In S. Gaglia & M.-O. Hinzelin (Eds.), *Inflection and word formation in Romance languages*. Oxford: Oxford University Press.

- Maiden, M. (2017). Cu privire la geneza vocalelor tematice în sistemul verbal dacoromanic. In A. Dragomirescu, A. Nicolae, C. Stan, & R. Zafiu (Eds.), *Sintaxa ca mod de a fi. Omagiu Gabrielei Pană Dindelegan, la aniversare* (pp. 247–54). Bucharest: Editura Universității din București.
- Maiden, M. (2018). *The Romance verb: Morphomic structure and diachrony*. Oxford: Oxford University Press.
- Maiden, M., Dragomirescu, A., Dindelegan, G. P., & Zafiu, R. (2020). *The Oxford history of Romanian morphology*. Oxford: Oxford University Press.
- Marantz, A. (1997). *Stem suppletion, or the arbitrariness of the sign*. Paper presented at the Université de Paris.
- McCarthy, J. J. (2005). Optimal paradigms. In L. J. Downing, T. Hall, & R. Faffelsiefen (Eds.), *Paradigms in phonological theory* (pp. 170–210). Oxford: Oxford University Press.
- McCarthy, J. J. & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Dickey, & S. Urbanczyk (Eds.), *University of Massachusetts Occasional Papers 18: Papers in Optimality Theory* (pp. 249–384). Amherst: University of Massachusetts, Graduate Linguistic Student Association.
- Mel’čuk, I. (1971). O suppletivizme. In S. K. Shaumjan (Ed.), *Problemy strukturnoj lingvistiki* (pp. 396–438). Moscow: Akademija nauk SSSR.
- Mel’čuk, I. (1976). On suppletion. *Linguistics*, 14(170), 45–90.
- Mel’čuk, I. (1994). Suppletion: Toward a logical analysis of the concept. *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”*, 18(2), 339–410.

- Mel'čuk, I. (2006). *Aspects of the Theory of Morphology*. Berlin: De Gruyter Mouton.
- Meul, C. (2010). The intra-paradigmatic distribution of the infix -I/ESC- from Latin to Modern Romance: morphomic patterning and beyond. *Morphology*, 20(1), 1–40.
- Meul, C. (2016). *The Romance reflexes of the Latin infixes -I/ESC- and -IDI-: restructuring and remodeling processes*. Hamburg: Helmut Buske Verlag.
- Meyer-Lübke, W. (1894). *Grammatik der Romanischen Sprachen: Bd. Formenlehre, vol. 2*. Leipzig: Reisland.
- Michael, K. (1996). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand & B. Laks (Eds.), *Current trends in phonology: Models and methods*, volume 1 (pp. 363–93). Manchester: University of Salford.
- Morlacchi, A. & Antelmi, D. (2005). L'interpretazione del linguaggio figurato nel ritardo mentale. *L'interpretazione del linguaggio figurato nel ritardo mentale*, (pp. 1000–1025).
- Mourin, L. (1977). Individualisation ou conformité dans les réfections des subjonctifs presents romans irréguliers. *Studii și cercetări lingvistice*.
- Narten, J. (1968). Zum "proterodynamischen" Wurzelpräsens. In J. C. H. et al. (Ed.), *Pratidānam: Indian, Iranian and Indo-European studies presented to F. B. J. Kuiper on his sixtieth birthday* (pp. 9–19). The Hague: De Gruyter Mouton.
- Nicoli, F. (1983). *Grammatica milanese*. Busto Arsizio: Bramante.
- Oltra-Massuet, I. & Arregi, K. (2005). Stress-by-structure in Spanish. *Linguistic Inquiry*, 36(1), 43–84.

- Oltra-Massuet, M. I. (2000). On the notion of theme vowel: a new approach to catalan verbal morphology. *MIT Occasional Papers in Linguistics*, 19.
- Osthoff, H. (1881). Zu der altlateinischen Duenos-Inschrift. *Rheinisches Museum für Philologie*, 36, 481–489.
- Osthoff, H. (1899). *Vom Suppletivwesen der indogermanischen Sprachen*. Heidelberg: J Hörning.
- Paul, H. (1886). *Principien der Sprachgeschichte*. Halle: Niemeyer, second edition.
- Penny, R. (2002). *A history of the Spanish language*. Cambridge: Cambridge University Press, second edition.
- Philippaki-Warbuton, I. (1993). The subjunctive mood and the syntactic status of the particle *na* in Modern Greek. *Folia Linguistica*, XXVIII(3-4), 287–326.
- Phillips, B. (2006). *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.
- Pizzuto, E. & Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of child language*, 19(3), 491–557.
- Pomino, N. & Remberger, E.-M. (2022). Does French have theme vowels? *Glossa: A Journal of General Linguistics*, 47(1).
- Rodríguez Bou, I. et al. (1952). *Recuento de vocabulario español*, volume 3. Río Piedras: Universidad de Puerto Rico.
- Rohlf, G. (1949). *Historische Grammatik der italienischen Sprache und ihrer Mundarten: Formenlehre und Syntax*, vol. 2. Bern: A. Francke.

- Rossini, G. (1975). *Capitoli di morfologia e sintassi del dialetto cremonese*. Florence: La Nuova Italia.
- Rudes, B. A. (1980). On the nature of verbal suppletion. *Linguistics*, 18(7-8), 655–676.
- Salerni, N., Assanelli, A., D’Odorico, L., & Rossi, G. (2007). Qualitative aspects of productive vocabulary at the 200-and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language*, 27(1), 75–87.
- Sandell, R. P. (2015). *Productivity in historical linguistics: Computational perspectives on word-formation in Ancient Greek and Sanskrit*. PhD thesis, Los Angeles: University of California, Los Angeles.
- Sanga, G. (1997). Lombardy. In M. Maiden & M. Parry (Eds.), *The dialects of Italy* (pp. 253–259). London: Routledge.
- Scheitlin, W. (1962). *Il pled puter: Grammatica ladina d’Engiadin’ota*. Samedan: Uniun dals Grischs.
- Sihler, A. (1995). *New comparative grammar of Latin and Greek*. Oxford: Oxford University Press.
- Smith, P. W., Moskal, B., Xu, T., Kang, J., & Bobaljik, J. D. (2019). Case and number suppletion in pronouns. *Natural Language & Linguistic Theory*, 37(3), 1029–1101.
- Sobolev, A. N. (2009). From synthetic to analytic case. In A. Malchukov & A. Spencer (Eds.), *The Oxford handbook of case* (pp. 716–729). Oxford: Oxford University Press.
- Spina, R. & Dressler, W. U. (2011). How far can diachronic change be predicted: The Italo-Romance first person plural present indicative. *Diachronica*, 28(4), 499–544.

- Starke, M. (2010). Nanosyntax: A short primer to a new approach to language. *Nordlyd*, 36(1), 1–6.
- Steriade, D. (2016). The morpheme vs. similarity-based syncretism. In A. R. Luís & R. Bermúdez-Otero (Eds.), *The Morpheme debate* (pp. 112–171). Oxford: Oxford University Press.
- Stump, G. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Stump, G. (2015). *Inflectional paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
- Stump, G. (2016). Morphomic categories and the realization of morphosyntactic properties. In A. R. Luís & R. Bermúdez-Otero (Eds.), *The Morpheme debate*. Oxford: Oxford University Press.
- Sturtevant, E. (1916). Dissimilative writing in Republican Latin and *UO* in Plautus. *Classical Philology*, 11(2), 202–207.
- Taglietti, G. (2005). Il cremonese. In A. Rognoni (Ed.), *Grammatica dei dialetti della Lombardia*. Milan: Mondadori.
- Taylor, C. (2007). Maximising stems. In M. Kokkonidis (Ed.), *Proceedings of LingO 2007* (pp. 228–235). Oxford: University of Oxford.
- Tekavčić, P. (1980). *Grammatica storica dell'italiano*. Bologna: Il mulino.
- Tonelli, L., De Marco, A., Vollmann, R., & Dressler, W. (1998). Le prime fasi dell'acquisizione della morfologia: un confronto tra l'italiano e il tedesco. *Cordin, P.; Iliescu, M. & Siller-Runggaldier, H. (a cura di): Parallela*, 6, 281–301.

- Tonelli, L., Dressler, W. U., & Romano, R. (1995). Frühstufen des Erwerbs der italienischen Konjugation. *Suvremena Lingvistika*, 21, 3–15.
- Tonelli, L. & Fabris, M. (2005). L’acquisizione della flessione verbale esemplificazione di un metodo di ricerca. *Annali della Facoltà di Lingue e Letterature Straniere dell’Università degli Studi di Sassari*, (pp. 2–13).
- Trommer, J. (2016). A postsyntactic morphome cookbook. In D. Siddiqi & H. Harley (Eds.), *Morphological metatheory* (pp. 59–93). Amsterdam: John Benjamins.
- Ussishkin, A. & Wedel, A. (2002). Neighborhood density and the root-affix distinction. In *North East Linguistics Society*, volume 32(2) (pp. 539–550).
- Van der Auwera, J., Dobrushina, N., & Goussev, V. (2003). A semantic map for imperative-hortatives. In D. Willems, B. Defrancq, T. Coleman, & D. Noël (Eds.), *Contrastive analysis in language: Identifying linguistic units of comparison* (pp. 44–66). Basingstoke: Palgrave Macmillan.
- Veselinova, L. N. (2006). *Suppletion in verb paradigms: Bits and pieces of the puzzle*. Amsterdam: John Benjamins.
- Volterra, V. (1976). A few remarks on the use of the past participle in child language. *Italian linguistics*, 2, 149–157.
- Weiss, M. L. (2020). *Outline of the historical and comparative grammar of Latin*. Ann Arbor, Michigan: Beech Stave Press.
- Wheeler, B. I. (1887). *Analogy and the scope of its application in language*. Ithaca, NY: John Wilson and Son University Press.
- Williams, E. (1994). Remarks on lexical knowledge. *Lingua*, 92, 7–34.

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT press.

Zanetti, U. (2005). Il bergamasco. In A. Rognoni (Ed.), *Grammatica dei dialetti della Lombardia*. Milan: Mondadori.

Zompì, S. (2023). **ABA in multidimensional paradigms: A MAX/DEP-based account*. PhD thesis, Cambridge, MA: Massachusetts Institute of Technology.