

Conjugation class from Latin to Romance:
heteroclisis in diachrony and synchrony



Steven James Kaye
Magdalen College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

2015

Abstract

Conjugation Class from Latin to Romance: heteroclisis in diachrony and synchrony

This thesis investigates the origins and behaviour of the non-canonical morphological phenomenon of *heteroclisis* in the verb paradigms of Latin and the Romance languages. Heteroclisis is the coexistence, within a single paradigm, of forms which pattern according to different inflectional classes existing otherwise in the language: a heteroclite lexeme can thus be seen as ‘mixed’ or ‘undecided’ as to its inflectional identity. I begin by examining the development of the theoretical concept of heteroclisis and approaches to the idea of inflectional class in general, before situating heteroclisis in typological space in comparison with better-known instances of non-canonical morphology such as deponency and suppletion; heteroclisis exists at a different level of generalization from these, because its identification presupposes the existence of inflectional classes, themselves generalizations over the behaviour of individual lexemes. I also consider two recent theoretical treatments of the phenomenon and survey recent linguistic studies making use of the notion. I then look at the synchronic and diachronic behaviour of heteroclisis in Latin and Romance verbs: the great time depth of our attestations of these languages gives us the chance to witness the development of successive examples of heteroclisis, and their subsequent treatment within the morphological system, in the history of a single family. Focusing chiefly on data from Latin, Romanian and Romansh, I find that the principal (though not the only) source for new instances of heteroclisis in Latin/Romance lies in regular sound change, and find that speakers can treat these synchronically anomalous patterns as robust models of inflectional behaviour to be extended over the lexicon or brought into line with pre-existing types of paradigm-internal alternation. These findings concur with previous demonstrations that speakers make use of non-canonical phenomena as markers of the internal structure of inflectional paradigms.

Acknowledgements

Though completing a doctorate is often seen as a solitary exercise, I could not have written this thesis without many people's help and support, and it is a privilege to be able to go some way towards acknowledging them here. First and foremost I am delighted to have this chance to thank Martin Maiden, who supervised my doctoral work with insight, patience and rigour. He was generous with his time and with his expertise in the many areas my work came to involve; his efforts secured the opportunity for me to carry out archive research in Romania and to spend a year of my doctoral programme at Leiden University; and this thesis would never have seen the light of day without his invaluable guidance and unfailing encouragement. It has been a pleasure to work with him, and I owe him a huge debt of gratitude.

I have also benefited from the generosity of many other scholars in the course of my research. At Oxford I have had the good fortune to be able to discuss different aspects of my work with Mary Dalrymple, Anna Kibort, Louise Mycock, Oana Uță, Mark Aronoff, Ilya Yakubovich, Wolfgang de Melo, Philomen Probert, John Penney, and the late Anna Morpurgo Davies, and it is a matter of regret that I was not able to pursue all the lines of research which these discussions opened up to me. My host at the Leiden University Centre for Linguistics, Geert Booij, gave me a great deal of his time during my stay there, both to discuss morphological matters and to introduce me to the university and the country: my year in the Netherlands was all the more rewarding, academically and otherwise, for his hospitality and kindness. I am also grateful to Paul Kiparsky, Greg Stump, Petros Karatsareas, Michiel de Vaan, Leonid Kulikov, Herman Seldeslachts, and Vincent Martzloff for sharing with me their insights, and in some cases their unpublished work, on Italic linguistics and the nature of inflection and heterocclisis. Marius Sala and Bogdan Marinescu, of the Romanian Academy in Bucharest, and Eugen Pavel, of the Pușcariu Institute in Cluj-Napoca, kindly granted me access to unpublished material from the archives of the New Romanian Linguistic Atlas; my time in Cluj is memorable for the warm welcome I received from Bogdan Harhătă and his family, as well as for all his help in the archives themselves. Gilles Authier kindly invited me to the TypoULM seminar at ENS (Paris) to present an early version of the research that eventually became Chapter 4. Finally here I reserve special thanks for my examiners, Frans Plank and John Charles Smith: their careful attention and advice helped me to view several issues in a new light and have improved the finished product enormously.

It has been a pleasure to undertake this work under the auspices of the Faculty of Linguistics, Philology and Phonetics at Oxford, which made it possible for me to carry out undergraduate teaching alongside my doctoral study and to rub shoulders with researchers in grammatical theory, classical languages and Romance linguistics alike. I am grateful in particular to the Research Centre for Romance Linguistics and

to the associated Romance Linguistics Seminar, where I also presented my work in progress. I would like to thank the Head of Linguistics, Aditi Lahiri, for her guidance and wise counsel throughout my DPhil and for the opportunity she gave me to base my research and teaching in the faculty's Language & Brain Lab. And as anyone in Oxford linguistics will know, it would be impossible to omit Aditi's stellar right-hand women Kate Dobson and Jane Cunning, whose task has ultimately been to keep the whole show on the road: like everyone else in the faculty, I have benefited both from their organizational prowess all along and from their immediate help in critical moments.

This doctoral work would have been impossible without the substantial financial assistance I received from the AHRC to fund my study at Magdalen College (award reference AH/H020632/1), as well as Oxford's Scatcherd European scholarship and the Jenkins Memorial scholarship awarded by the Europaeum, which jointly funded my year of study at the Leiden University Centre for Linguistics: I gratefully acknowledge this generous support. It is a privilege to have had Magdalen College as my home from home for many years, and my thanks go to Toby Garfitt for his mentorship during my time at the college and to Susan Burton, Katie O'Connor and Nancy Cowell for their work dealing with the administrative demands involved in keeping me there. In Leiden, Gea Hakker ensured that my stay as a guest researcher at LUCL proceeded as smoothly as possible.

Throughout my time as a doctoral student, I have been lucky enough to enjoy the friendship of many fantastic people, in Oxford, Leiden and beyond. Whether students of linguistics or not, they have undoubtedly had to hear more about heteroclisia in the last few years than they could ever have hoped to, and for that I apologize here. I must take this opportunity to thank in particular Nikos Koutsoukos, Louise Esher, Paul O'Neill and Nick Zair for years of enlightening conversation on the issues treated in this thesis, and Kostis Dimos for navigating the recesses of the University of Zürich library system on my behalf; but to all my friends, thank you for the intellectual, practical and moral support you have shown me, which I could not have done without. Special mention must go to the many people I have shared the Lab with over the years, dating back to the time of my MPhil: we have been through a lot together. For all the advice, tea, nonsense, amazing cakes and last-minute rescues my thanks go to Holly, Jeannique, Ben, Samantha, Allison, Phil, Lara, Adam, Karola, Franziska, Colin, Cong, Johanneke, Hilary – and most of all Sandra, whose constant encouragement, generosity and help of all kinds were crucial in seeing me through.

Finally, I want to thank my family, to whom this thesis is dedicated, for their love and support.

Table of Contents

Index of tables	i
1. Introduction	1
1.1 Subject and aims	1
1.2 Structure of the present work	13
2. Heterocclisis and inflectional class in linguistic scholarship	15
2.1 The meaning of ‘heterocclisis’	15
2.2 Inflectional diversity and inflectional classes	30
2.2.1 <i>Carstairs-McCarthy and affix economy</i>	32
2.2.2 <i>Inflectional diversity in Natural Morphology</i>	45
2.2.3 <i>Discussion</i>	57
2.3 Heterocclisis in Canonical Typology	80
2.4 Noyer 2005 on heterocclisis and ‘systematic mixed inflection’	87
2.5 Heterocclisis in Paradigm Function Morphology (Stump 2006)	100
2.6 Three recent studies invoking heterocclisis	111
2.6.1 <i>Karatsareas 2011a on Asia Minor Greek</i>	111
2.6.2 <i>Luís 2008 on Caló</i>	115
2.6.3 <i>Esher 2012 on Occitan</i>	118
3. Heterocclisis in the Latin verb	128
3.1 Introduction	128
3.1.1 <i>The Latin verb paradigm and the (present-stem) conjugations</i>	129
3.1.2 <i>An overview of the PIE verb</i>	135
3.1.3 <i>From PIE to the Latin conjugations</i>	144
3.2 The <i>capiō</i> type	150
3.2.1 <i>Morphological properties of the capiō type</i>	150
3.2.2 <i>The origins of heterocclisis in the capiō type</i>	157
3.2.3 <i>The prehistory of the capiō type and its relationship with the fourth conjugation</i>	163
3.2.4 <i>Conclusion</i>	172

3.3	The <i>domō</i> type	175
3.3.1	<i>Further background on the Latin verb system</i>	180
3.3.2	<i>The origins of the domō type</i>	187
3.3.3	<i>Conclusion</i>	196
3.4	The <i>petō</i> type	198
4.	Romanian	213
4.1	Introduction and overview of the Romanian verb	213
4.2	Heteroclisis as the indirect result of sound change: Maiden 2009	223
4.3	Beyond <i>a coase</i> and <i>a țese</i>	234
4.4	Standard Romanian phenomena	246
4.5	Conclusion	261
5.	Heteroclisis in Romansh	264
5.1	Introduction	264
5.2	Pan-Romansh heteroclisis: fourth conjugation morphology ‘out of place’	267
5.3	From heteroclisis to new regularity? The case of Surselvan	282
5.4	The infinitive in Val Müstair	291
6.	Conclusions	311
	Bibliography	328

Index of Tables

Chapter 1

Table 1.1	Heteroclite inflection of <i>domus</i> ‘house’ in Classical Latin	2
Table 1.2	Heteroclite inflection of <i>pramen</i> ‘spring’ in Standard Czech (Stump 2006: 280, slightly modified; Heim 1982: 22, 41f., 176)	4

Chapter 2

Table 2.1	Semi-deponent inflection of Latin <i>fīdō</i> ‘trust’ (cf. Allen and Greenough 1903: §192)	23
Table 2.2	Heterocclisis and deponency in Gothic preterite-presents (Baerman 2007: 17)	26
Table 2.3	Heterocclisis and stem alternation in Romanian <i>a fi</i> ‘be’ (Gönczöl-Davies 2007: 96-101)	27
Table 2.4	Verbal inflectional classes in Burmeso (Corbett 2009a: 9; Donohue 2001: 100, 102)	34
Table 2.5	Inflectional suffixes of the Russian noun (after Brown and Hippisley 2012)	35
Table 2.6	Inflectional behaviour of German <i>Dorn</i> ‘thorn’ alongside representatives of ‘strong’ and ‘weak’ macroparadigms (after Carstairs 1987: 241-50)	39
Table 2.7	A pattern outlawed by the NBP (Enger 2007: 283)	40
Table 2.8	A breach of the Paradigm Economy Principle in Latin nouns (after Carstairs 1984)	43
Table 2.9	A partial paradigm of some Latin ‘third declension’ nouns (after Wurzel 1984: 120)	50
Table 2.10	Distinctive characteristics of strong inflection in Standard German (after Bittner 1996: 55-61, Dammell 2011: 196-98)	54
Table 2.11	The properties of canonical inflection (after Corbett 2009a: 2)	82
Table 2.12	Schema for Systematic Mixed Inflection (modified slightly from Noyer 2005: 274)	92
Table 2.13	Second schema for Systematic Mixed Inflection (modified slightly from Noyer 2005: 277)	95
Table 2.14	Mixed inflection in Gothic masculine <i>i</i> -stem nouns (modified from Noyer 2005: 275)	97
Table 2.15	Three nominal paradigms belonging to different declensions in Standard Modern Greek	111

Chapter 3

Table 3.1	Outline of the paradigm of <i>moneō</i> ‘warn’	130
Table 3.2	Some <i>infectum</i> forms of first conjugation <i>cūrō -āre</i> ‘care for’ and third conjugation <i>regō -ere</i> ‘rule’	133
Table 3.3	<i>Infectum</i> forms of <i>capiō</i> ‘take’	151
Table 3.4	<i>Infectum</i> forms of (third conjugation) <i>regō</i> ‘rule’	151
Table 3.5	<i>Infectum</i> forms of (fourth conjugation) <i>sentio</i> ‘feel’	151
Table 3.6	Present, perfect and third stem forms of verbs of the <i>domō</i> type	176
Table 3.7	Inflected forms of heteroclite <i>domō</i> ‘tame’, in comparison with corresponding first and non-first conjugation forms	177
Table 3.8	Present, perfect and third stem forms of verbs of the <i>petō</i> type	200

Chapter 4

Table 4.1	First conjugation (without augment), <i>a cânta</i> ‘sing’	216
Table 4.2	Second conjugation, <i>a putea</i> ‘be able’	216
Table 4.3	Third conjugation (u-type), <i>a țese</i> ‘weave’	216
Table 4.4	Third conjugation (s-type), <i>a arde</i> ‘burn’	217
Table 4.5	Fourth conjugation (without augment), <i>a dormi</i> ‘sleep’	217
Table 4.6	Fourth conjugation (with augment), <i>a citi</i> ‘read’	217
Table 4.7	The paradigm of <i>a coase</i> ‘sew’ in standard Romanian	223
Table 4.8	The paradigm of <i>a coase</i> in Dobrița	225
Table 4.9	The paradigm of <i>a coase</i> in Beloțu (after Maiden 2009: 66f.)	229
Table 4.10	The paradigm of <i>a sughita</i> ‘hiccup’ in standard Romanian	234
Table 4.11	The paradigm of <i>a sughita</i> in Plavișevița (Banat)	236
Table 4.12	The paradigm of <i>a sughita</i> in Cornereva (Banat)	237
Table 4.13	Partial paradigm of <i>saliō</i> ‘jump’ in Latin, and postulated forms after the development of yod	240
Table 4.14	Illustration of the L-pattern (postulated) after fusion of the cluster /l + j/	240
Table 4.15	Partial paradigm of <i>a sări</i> illustrating stem alternation in N-pattern cells	242
Table 4.16	Partial paradigm of <i>a zbură</i> , in the meaning ‘jump’, in Bălan (Crișana)	243
Table 4.17	Partial paradigms of <i>a (a)mâna</i> and <i>a rămânea</i>	244
Table 4.18	The unaugmented ‘fifth conjugation’ in <i>î</i>	250

Table 4.19	The paradigm of <i>a acoperi</i> ‘cover’ in standard Romanian	252
Table 4.20	Schema for systematic mixed inflection (cf. Tables 2.12 and 2.13)	257
Table 4.21	The Romanian conjugations and <i>a acoperi</i>	258
Chapter 5		
Table 5.1	Paradigm of a regular (‘alternating’) first conjugation verb	268
Table 5.2	Paradigm of a regular third conjugation verb	269
Table 5.3	Paradigm of a regular (‘augmented’) fourth conjugation verb	269
Table 5.4	The paradigm of the heteroclite verb <i>cuorrer</i> ‘run’ in Vallader	274
Table 5.5	Numbers of Surselvan speakers indicating a preference for each of the two available 1pl present indicative endings on selected verbs (Winzap 1990: 40, with modifications)	286
Table 5.6	Val Müstair infinitives	293

1. Introduction

1.1 Subject and aims

This thesis investigates the morphological phenomenon of *heteroclisis* in the history of the Latin and Romance verb. Bringing together case studies treating instances of heteroclisis in Latin and Romance languages, it aims both to provide a descriptive survey of the range of heteroclite behaviour identifiable over the historical development of the family – the first such investigation of heteroclisis dedicated to a coherent linguistic domain in diachrony – and to show the variety of ways in which this behaviour has emerged and undergone further evolution against the backdrop of the inflectional system as a whole. Morphologists of different stripes have recently shown an increased interest in the concept of heteroclisis, but relatively little descriptive and/or empirical work has been carried out on the topic in comparison with other morphological phenomena of a similar nature. At the same time, the relevance of heteroclisis to a satisfactory description of the verb systems of Latin and its daughter languages, and the extent to which it is tightly bound up with the development of better-known characteristics of these systems, has generally been underappreciated. The present work is thus intended as a contribution both to the general morphological literature and to the historical linguistics of Italic and Romance.

Although we will see that other meanings have been attested for the term, and authors still differ in detail with respect to the scope they give to it, the core meaning of *heteroclisis* as it is used in most modern morphological scholarship (including this thesis) concerns the relationship between lexemes and inflectional classes within a language's synchronic system, where inflectional class can be assigned a definition

along the lines of Aronoff's (1994: 64): 'An inflectional class is a set of lexemes whose members each select the same set of inflectional realizations'. What marks out the paradigm of a heteroclite lexeme, or group of lexemes, is that it does not pattern exclusively with any single one of the inflectional classes which can otherwise be identified for the language – but instead contains some cells which show morphological behaviour proper to one of these classes, alongside other cells which show morphological behaviour only proper to others. In an instance of heteroclitism, a single paradigm thus seems to be 'split' between two or more established inflectional classes.¹

Table 1.1 Heteroclite inflection of *domus* 'house' in Classical Latin

	2 nd declension <i>servus</i> 'slave'		heteroclite <i>domus</i> 'house'		4 th declension <i>manus</i> 'hand'	
	sg	pl	sg	pl	sg	pl
nom	<i>servus</i>	<i>servī</i>	<i>domus</i>	<i>domūs</i>	<i>manus</i>	<i>manūs</i>
acc	<i>servum</i>	<i>servōs</i>	<i>domum</i>	<u><i>domōs</i></u>	<i>manum</i>	<i>manūs</i>
gen	<i>servī</i>	<i>servōrum</i>	<i>domūs</i>	<u><i>domōrum</i></u>	<i>manūs</i>	<i>manuum</i>
dat	<i>servō</i>	<i>servīs</i>	<i>domuī</i>	<i>domibus</i>	<i>manuī</i>	<i>manibus</i>
abl	<i>servō</i>	<i>servīs</i>	<u><i>domō</i></u>	<i>domibus</i>	<i>manū</i>	<i>manibus</i>

The basic phenomenon is illustrated in Table 1.1, which gives the inflectional paradigm of the noun *domus* 'house' in Classical Latin (Meiser 1998: 150) alongside those of *servus* 'slave' and *manus* 'hand', regularly inflecting nouns of the second and fourth declensions respectively. As the table shows, certain cells of the paradigm of *domus* contain forms (underlined in the table) which pattern distinctively in accordance with the second declension and unlike the corresponding forms seen in the fourth: ablative singular *domō* and accusative and genitive plural *domōs*, *domōrum*. At the same time, however, other cells contain forms (in bold in the table)

¹ In light of Aronoff's definition, statements such as this may be thought to beg the question, in that *domus* could alternatively be said to belong to an inflectional class showing the behaviour seen here and containing precisely one member. This point will be pursued in the following chapter (especially section 2.2.3), which provides a more general discussion of the nature of inflectional diversity and the idea of inflectional classes.

which pattern distinctively in accordance with the fourth declension and unlike those seen in the second: genitive and dative singular *domūs*, *domuī* and nominative, dative and ablative plural *domūs*, *domibus*, *domibus*. The coexistence in this single paradigm of inflectional behaviour belonging to these two patterns, which usually stand opposed to each other in that they characterize separate sets of nouns, makes *domus* an example of a heteroclite lexeme.

As it happens, there are also two further inflected forms in the paradigm of *domus* (the nominative and accusative singular *domus*, *domum*) which cannot be identified with either the second or the fourth declension in particular. As a result, the paradigm of *domus* cannot be divided neatly into second declension and fourth declension zones lying alongside each other without any overlap. However, this is immaterial to an identification of *domus* as heteroclite. There is no requirement that a heteroclite paradigm should be made up of entirely disjoint sets of cells each affiliated with its own inflectional class. What makes *domus* heteroclite is the morphological ‘incompatibility’ between some of its inflected forms and others, e.g. gen sg *domūs* and gen pl *domōrum*, when these are viewed in the light of the nominal inflectional system as a whole.

The inflectional behaviour of Classical Latin *domus* is unique to that lexeme: no other noun shows the same morphological opposition between second and fourth declension forms in precisely these arrays of cells. Lexical uniqueness is not a defining feature of heteroclis, however. Stump (2006: 279f.) introduces the notion of heteroclis with the example of the Czech masculine noun *pramen* ‘spring, source’, whose paradigm (as shown in Table 1.2) makes use of inflectional endings belonging to two opposing inflectional classes, labelled by Stump as ‘soft-masculine’ and ‘hard-masculine’: where these two declensions disagree, every case form of

pramen in the singular shows distinctively soft-masculine inflection (underlined), while every case form in the plural shows distinctively hard-masculine inflection (in bold). This unusual pattern of behaviour is in fact not restricted to *pramen*: it is shared by several other masculine nouns in Czech, including *kámen* ‘stone’ and *řemen* ‘strap’ (Heim 1982). Nonetheless, although they form a coherent group within the Czech nominal lexicon, we have the same grounds for identifying heteroclisis in these nouns as we have in the case of Latin *domus*: certain case forms in the paradigm of *pramen* clearly ‘belong’ to a pattern of inflectional behaviour which, equally clearly, does not characterize other case forms of the same noun.

Table 1.2 Heteroclite inflection of *pramen* ‘spring’ in Standard Czech (Stump 2006: 280, slightly modified; Heim 1982: 22, 41f., 176).

	soft-masculine <i>pokoj</i> ‘room’		heteroclite <i>pramen</i> ‘spring’		hard-masculine <i>most</i> ‘bridge’	
	sg	pl	sg	pl	sg	pl
nom	<i>pokoj</i>	<i>pokoje</i>	<i>pramen</i>	<i>prameny</i>	<i>most</i>	<i>mosty</i>
voc	<i>pokoji</i>	<i>pokoje</i>	<u><i>prameni</i></u>	<i>prameny</i>	<i>moste</i>	<i>mosty</i>
acc	<i>pokoj</i>	<i>pokoje</i>	<i>pramen</i>	<i>prameny</i>	<i>most</i>	<i>mosty</i>
gen	<i>pokoje</i>	<i>pokojů</i>	<u><i>pramene</i></u>	<i>pramenů</i>	<i>mostu</i>	<i>mostů</i>
dat	<i>pokoji</i>	<i>pokojům</i>	<u><i>prameni</i></u>	<i>pramenům</i>	<i>mostu</i>	<i>mostům</i>
instr	<i>pokojem</i>	<i>pokoji</i>	<i>pramenem</i>	<i>prameny</i>	<i>mostem</i>	<i>mosty</i>
loc	<i>pokoji</i>	<i>pokojích</i>	<u><i>prameni</i></u>	<i>pramenech</i>	<i>mostě</i>	<i>mostech</i>

The phenomenon of heteroclisis thus represents a type of inflectional exceptionality at the level of the lexeme, but one based in regularities identifiable at the level of the lexicon. Individual heteroclite lexemes, or groups of lexemes, have inflected forms which show no sign of irregularity when considered in isolation: it is only when taken together that they go against generalizations which can otherwise be made about the inflectional system as a whole. The shape of the genitive singular form of *domus* gives rise to an expectation about the shape of the genitive plural form which is not fulfilled, and vice versa, but neither (fourth declension) genitive singular *domūs* nor (second declension) genitive plural *domōrum* can be considered irregular in its own right.

The fact that a lexeme is heteroclit also brings with it no implication that it is irregular in any other way beyond its morphological behaviour. Inflected forms belonging to a heteroclit paradigm operate in their syntactic context exactly as would be expected of any inflected forms realizing the given morphosyntactic features: the formal relationship between the ways these features are realized in different cells of the paradigm is of no syntactic consequence. Like the existence of inflectional classes themselves, heteroclysis is fundamentally a morphological fact, with no correlate in the syntax and no direct relevance outside the morphological system.

This central characteristic of heteroclysis is one reason why the concept has seen something of a rise in interest among morphological theorists in recent years, as the idea that a language's morphology is a linguistic domain in its own right, and not merely an epiphenomenon of the syntactic and phonological facts of that language, has gradually regained general (though still far from unanimous) acceptance in reaction to the 'syntactocentric' (Jackendoff 2002) view which initially prevailed in linguistics following the rise of generative grammar. The existence of heteroclysis, like the existence of inflectional classes themselves, provides support for the claim that morphology is at least partially autonomous and can operate according to principles independent from those required in other domains. Accordingly, the nature of heteroclysis has been touched on (at least) by researchers interested in highlighting and charting these autonomous principles, and potentially providing a theoretical model for them. Meanwhile, those who ultimately do not wish to grant any independent theoretical status to morphology are nonetheless forced to recognize in at least some instances of heteroclysis (in the sense used here) a inflectional

phenomenon whose existence needs to be taken seriously, and thus requires a principled account in keeping with the overall model of grammar they propose.

Heterocclisis is far from the only example of a grammatical phenomenon peculiar to the morphological system: others include deponency and (as mentioned above) the existence of inflectional classes, which are a prerequisite for the notion of heterocclisis. A recent approach to morphological analysis brings out a feature shared by all these (and other) concepts used in the description of inflectional systems: namely, that they all represent deviations from the theoretically ‘ideal’, or *canonical*, situation in which every distinction that the inflectional morphology is called on to encode in a given language is encoded consistently, efficiently and unambiguously by the language’s morphological resources. It is immaterial whether this canonical state of affairs can be found in (any coherent part of) the inflectional morphology of a genuine human language: rather, the *canonical typology* approach (Corbett 2007a) is designed to draw attention to the large number of dimensions along which attested inflectional behaviour can diverge from this ideal, and thus provide an abstract ‘typological space’ in which particular phenomena studied by morphologists can be situated. This approach thus raises the question of exactly where heterocclisis ‘fits’ in an abstract descriptive system of this kind, in relation to other types of non-canonical morphological behaviour: this is a logical question, not specific to any particular language.

At the same time, however, there are more concrete questions to be asked about heterocclisis as it is found in particular cases. What patterns of heterocclisis occur, and do they show any consistent tendencies within a given language (or across languages)? What role can heterocclisis play in the synchronic inflectional system? Closely related, and equally crucial, are diachronic questions. How do the various

instances of heteroclisis arise, and what further developments may they undergo once they have taken root in a given language?

The particular relevance of exceptionality and non-canonical behaviour to the description of morphological systems, as highlighted in Simon and Wiese (2011: 20-23), is reflected in the practice of typologists concerned with morphology: for example, in the last decade Matthew Baerman and his colleagues at the Surrey Morphology Group have successively produced a monograph on syncretism (Baerman, Brown and Corbett 2005) and overseen edited volumes dedicated to deponency (Baerman et al. 2007) and defectiveness (Baerman, Corbett and Brown 2010). The absence of any similarly extensive treatment of heteroclisis, under the auspices of the SMG or elsewhere, reflects its relatively low profile in morphological research in comparison with these better-established phenomena. Stump 2006, an article offering an account of heteroclisis in terms of Stump's own theory of Paradigm Function Morphology, at the same time represents the most substantial cross-linguistic survey of heteroclisis seen to date, listing a wide range of examples of the phenomenon from inflectionally rich language families around the world (as its programme is partly to demonstrate that heteroclisis is not a rare phenomenon limited to Indo-European languages). However, the bulk of this article is concerned with the formal analysis of specific existing patterns involving heteroclisis in terms of a general theory of 'paradigm linkage' (an important mechanism made available by PFM), and no weight is laid on the question of how a given pattern of heteroclisis may have arisen against the background of the whole inflectional system relevant to lexemes of a particular syntactic category, or how the emergence of a given instance of heteroclisis may in turn have had ramifications for this system.

The present study takes a different and in some ways complementary approach. I look at the emergence and development of several prominent instances of heteroclisis as they form part of the evolution of a single inflectional system over the long term – that of the Latin/Romance verb. The focus is thus on the ways in which both the origins and the synchronic characteristics of these instances of heteroclisis may be bound in with more general diachronic trends and synchronic patterns. Taken together, the case studies from Latin and Romance presented here form a partial history of the inflectional morphology of the verb viewed with the phenomenon of heteroclisis at its centre rather than at the margins.

The Latin/Romance family, and specifically its verb system, is highly suitable as the subject of a historical study of this kind for a number of reasons. It is unusually rich among language families in that we both know a great deal about the history and structure of its many contemporary representatives, including numerous standard languages with a long history of written attestation, and also possess copious records of (something very close to) the ancient parent language from which all these varieties descend. Meanwhile, Latin itself is not known to us in isolation but is flanked by Italic sister languages of similar antiquity as well as more distantly related languages belonging to other branches of Indo-European, both of which provide evidence allowing us to extend our understanding of its structural development further back into the past. The morphological system of the verb employs a largely stable set of inflectional classes which can be identified across the daughter languages and can be traced back as far as the common Italic period. What is more, the synchronic and diachronic behaviour of the verb paradigm in Latin and the Romance languages has been the subject of much of the research carried out in the ‘autonomous morphology’ tradition and has provided some of the clearest

evidence for purely morphological paradigmatic effects (cf. in particular the work on Romance ‘morphomic’ patterns outlined towards the end of section 2.5).

While Latin barely features,² and Romance not at all, among the sources of heteroclisis mentioned in Stump 2006, in the intervening years since this article helped to bring the phenomenon to widespread linguistic attention, a number of studies have appeared which label as heteroclisis the behaviour of particular (large or small) groups of verbs in several of these languages. As will be seen, in many of the sections in the present thesis I consider these identifications, and as a rule endorse them (however, I reject one recent suggestion that heteroclisis is a more appropriate concept than deponency for capturing the nature of what have traditionally been called the Latin ‘deponent verbs’ (see section 3.1, fn. 49)). At the same time, even in this well-studied family, the prevalence of heteroclisis and its importance in the development of the inflectional system of the verb have been insufficiently appreciated. In bringing together and treating in more detail the various claims made so far for heteroclisis in Latin/Romance, as well as pointing to many further instances of the phenomenon which have not previously been discussed in these terms, I aim to provide a more representative picture.

This thesis is thus intended not only as a contribution to the study of heteroclisis but also as a contribution to the historical linguistics of Italic and Romance. Naturally, it does not attempt to document every instance of heteroclisis in the history of the Italic/Romance verb. Rather, as my interest lies primarily in the ways in which heteroclisis fits within the inflectional system as a whole and may reflect its existing characteristics or give rise to new ones, I devote little attention to truly isolated instances (whose status is equivalent to that of *domus* in the nominal

² Three Latin lexemes (one of them *domus*) are mentioned as showing heteroclisis, but only in passing: no account is given of their inflectional behaviour or how it qualifies as heteroclitite.

system) and concentrate on those cases which show some measure of systematicity and affect groups of lexemes in a coherent way – although, as will be seen, no dividing line can be drawn between ‘merely lexical’ instances of heteroclisis and those which have wider ramifications for the inflectional system. I also focus particularly on certain varieties within Romance (namely Romanian and the Romansh dialects) whose low profile in comparison to better-known and more widely spoken sister languages means that their morphological history has remained relatively underexplored.

At the same time, in considering the origins and behaviour of heteroclisis in Latin together with these Romance varieties, I wish to bring out the essential unity and continuity of a verb system whose study is, for the most part, still divided up between researchers belonging to two different scholarly traditions. It is understandable that work targeted towards Romanists generally treats the Latin of the classical period as a *fait accompli*, since it represents (at least a close approximation of) the starting point for Romance developments; but this disguises the extent to which we also understand the earlier historical and prehistoric developments which led to the Latin system itself and can thus appreciate it as an intermediate stage in a larger process. I aim to present something of this longer view here, following the approach taken in other recent studies bringing together traditionally Latinist and Romanist scholarship in particular domains, such as Sampson 1999, 2010 on phonological topics, Ledgeway 2012 on morphosyntax, and Adams 2007, 2013 on aspects of sociolinguistic history.

As this study is interested in heteroclisis in terms of both the patterns it shows and its development over time, by pursuing these two topics in parallel I also aim to highlight a systematic ambiguity which has arisen in the use of the term *heteroclisis*

in the morphological literature. Introducing the concept of heterocclisis above, I treated it as a label relevant to the *synchronic* description of inflectional behaviour. However, the term has not only been used to mark out lexemes showing specific patterns of inflectional morphology in the context of a given synchronic system: it is also used on occasion as a characterization of changes in inflectional class affiliation in part of the paradigm of certain words, i.e. as a label for what seems intuitively to be the corresponding *diachronic* phenomenon. This extension of meaning is a natural one. However, the notions of ‘synchronic’ and ‘diachronic’ heterocclisis are logically distinct: this lack of a necessary connection between the two is illustrated by several of the case studies presented here. One of the main general aims of this thesis is to use the particular Latin/Romance examples found here to demonstrate the importance of recognizing this distinction.

This ambiguity, which reflects the conflation of synchronic and diachronic phenomena, is orthogonal to other disagreements which are found regarding the meaning of heterocclisis. The term has a long history as a label for certain kinds of unusual inflectional behaviour, but the use standard in modern morphological theory, in which it refers to the phenomenon treated here and involving the relationship between lexemes and inflectional classes, has not always been paramount. Notably, in the philological tradition it has been most consistently applied to nouns belonging to a particular historical type (the ‘*r/n*-stems’) which is reflected widely across Indo-European and reconstructed for the proto-language, without implying any particular view about the inflectional behaviour of those nouns in a given daughter variety. While the present thesis is not primarily about the term ‘heterocclisis’, then, but about the inflectional phenomenon introduced under that name above as it is found in the development of the Latin/Romance verb system, it is important to sketch out

something of the wider range of meanings which 'heterocclisis' has held in linguistic scholarship.

1.2 Structure of the present work

This study is thus structured as follows. In **Chapter 2**, I begin by briefly discussing the history of ‘heteroclisis’ and the different uses to which this term has been put, before moving on to an extensive investigation of inflectional diversity and the ways in which different scholars have taken it to be constrained: this involves a discussion of the notion of inflectional classes, upon which heteroclisis depends. Introducing the Canonical Typology framework, I consider the position of heteroclisis in typological space in relation to better-known types of non-canonical inflectional behaviour, and point to some of the properties which would in turn characterize ‘canonical heteroclisis’. I then present the two most prominent recent contributions to a theoretical understanding of heteroclisis, the second of which highlights the importance of ‘purely morphological’ patterns within the paradigm, as have been the subject of intense study in Romance historical linguistics. Finally, I turn to three descriptions of particular language-specific phenomena in terms of heteroclisis, pointing out the potential interest afforded by the detailed historical study of heteroclisis as well as the value of a clear-sighted approach to the distinction between synchrony and diachrony in this area.

Chapter 3 concerns instances of heteroclisis in Latin. In section 3.1, I give an overview of the inflectional system of the Latin verb and the division of the verb lexicon into conjugations, before giving an account of the Indo-European origins of this general system and the developments which led to the shape it takes in Latin. Sections 3.2 to 3.4 then each treat one group of heteroclite verbs, showing to what extent their behaviour can be explained as the result of the same forces responsible for producing the regular patterns they stand out against: phonological,

morphological and semantic factors are all shown to have had a hand in the emergence of the different forms of heteroclisis observed.

The remainder of the thesis discusses heteroclisis in the verb systems of Romance. I turn first to Romanian (**Chapter 4**), which in the inflection of certain verbs shows patterns of heteroclisis dialectally that have already figured (Maiden 2009) in the theoretical debate as to how the phenomenon in general should be modelled. I briefly recapitulate the synchronic and diachronic situation, explain the relevance of these particular Romanian facts for this question and weigh up the merits of Maiden's 'purely morphological' approach. I also examine the implications of heteroclisis in the Romanian verb for the proposal that some instances of heteroclisis (in the terms used here) are created by *interclass syncretism* (Noyer 2005), as explained below in section 2.4. Next, **Chapter 5** looks at heteroclisis in the dialects of Romansh. One prevalent pattern, treated in sections 5.2 and 5.3, resembles that discussed by Maiden for Romanian: however, I show that despite this similarity the Romansh pattern is unlikely to have arisen for the same reason. Moreover, there is some evidence of a change in progress in certain Romansh dialects, in which this originally heteroclitic behaviour has spread across the lexicon to the extent that it can now be considered to characterize a conjugation in its own right: this again brings out the importance of the distinction between synchrony and diachrony stressed in the present work. I also investigate (section 5.4) a separate diachronic phenomenon restricted to a single valley located at the very edge of the Romansh-speaking area: here, what can be identified as a process of conjugational incursion in diachrony is apparently connected with a general recasting of the formal relationship between the infinitive and certain finite forms of the paradigm. **Chapter 6** draws together my conclusions.

2. Heterocclisis and inflectional class in linguistic scholarship

2.1 The meaning of ‘heterocclisis’

As I have introduced the term here and will use it in the course of this thesis, the notion of *heterocclisis* concerns the relationship between lexemes and inflectional classes: the paradigm of a heterocclite lexeme brings together types of morphological behaviour which can be recognized as characterizing distinct classes in the language in the light of its inflectional system as a whole. As will be seen, something akin to this meaning is intended by most uses of the term in current morphological scholarship specifically addressing inflectional systems, although researchers do differ in the precise scope they grant to it.

However, this restricted conception of heterocclisis as being concerned with the juxtaposition of different inflectional classes in particular has only ‘crystallized’ recently from a state of affairs in which, across linguistic scholarship in general, the meaning of the term was potentially somewhat broader in one sense but at the same time narrower in another. This is because ‘heterocclisis’ has traditionally been a label assigned almost exclusively to the behaviour of nouns; meanwhile, although the possibility of using it in reference to phenomena involving inflectional class is not an invention of recent scholarship, it has equally been used to refer more generally to nouns which show any kind of irregularity in the paradigm, most usually as concerns stem shape as opposed to choice of endings. Examples of this sense of ‘heterocclisis’ may be found in definitions from standard handbooks, cf. Conrad 1985 (s.v. *Heteroklitikon*): ‘Substantiv, das mehrere Wortstämme bei der Kasus- und

Numerusbildung verwendet; z.B. russ. *mat'* – *materi* ‘Mutter’.³ The formulation still given in Matthews 2007 (s.v. *heteroclite*) is in keeping with this broad usage, and indeed the intended scope could hardly be broader: ‘a noun whose inflection follows something other than the regular pattern is traditionally “heteroclite”’.⁴

This very general definition of the term ‘heterocclisis’, encompassing unusual nominal behaviour not necessarily linked with variability in inflectional class, is inherited from the practice of the classical grammarians: cf. Schad (2007: 195f.), who finds that in Latin grammatical works *heterocclitus* (following Greek *heteróklitos* ‘inflected otherwise’) can refer to nouns showing stem suppletion or irregularity of various kinds. The potential of the classical term to point to abnormalities in stem shape, among other things, lies behind what has been its predominant use in Indo-European philology, where at least one type of nominal declension traditionally labelled ‘heteroclite’ is reconstructed for the proto-language – namely the *r/n*-stems (Fortson 2004: 110f.), surviving e.g. in Hittite *wātār* gen. *witen-aš* ‘water’ and Latin *femur* gen. *femin-is* ‘thigh’, which seem to have shown stem allomorphy across the paradigm that is not predictable on the basis of any principle of Indo-European morphophonology. The label is not intended to presuppose that we can also posit regular *r*- and *n*-declensions for the proto-language (or the relevant daughter languages), whose morphology these nouns use and against which they stand out.

Classic treatments of this phenomenon as a puzzling feature of nominal morphology in the Indo-European proto-language (and one potentially reflecting traces of the regular behaviour of a still earlier system) include Petersson 1921,

³ Significantly, Russian *mat'* ‘mother’, whose genitive is *materi* rather than **mati* as would be expected on the basis of the nominative, is not heteroclite in the sense relevant to this thesis, as it does not pattern with distinct declensions over the paradigm: indeed it is mentioned by Stump (2006: 283) precisely as an example of a noun which shows irregular stem suppletion without thereby showing heterocclisis.

⁴ This is in essence identical to the earliest definition of *heterocclite* in its grammatical use cited in the OED, from Blount’s *Glossographia* (1656): ‘*Heterocclite*, that is declined other than common Nouns are.’

Benveniste 1935 and Schindler 1975, while its remnants in Hittite – where the *r/n*-stem declension remains productive – and Latin are treated respectively with a view to their Indo-European origins by Kammenhuber 1955 and Rix 1965. The salient feature of ‘heteroclis’ in this sense, apart from the interest and difficulty it poses for both comparative and internal reconstruction, is the characteristic distribution taken up by the opposing stems within the paradigm, whereby the stem found in the nominative (generally = accusative) singular differs from that found elsewhere: note that the identity of the nominative and accusative forms, at least in the uniformly neuter *r/n*-stems, is unremarkable as nominative and accusative always pattern together in Indo-European neuters (Giacalone Ramat and Ramat 1998: 83).

The exceptional nature of such behaviour in the context of ancient Indo-European noun inflection has garnered it a great deal of interest, with the result that within classical philology the *r/n*-stems and nouns showing similar behaviour are treated as *the heteroclitites par excellence*, the term often being employed exclusively for nouns of this kind even where other forms of phonologically unmotivated stem alternation do exist in a given language. Owing to the traditional prominence accorded to Indo-European within the wider linguistic sphere, this conception of heteroclis is also the one which has had most currency outside the study of Hittite, Sanskrit, Latin and other languages where the inherited Indo-European formations themselves survive. Thus despite the very general definitions of heteroclis which can be observed in the manuals just cited, it can be fairly stated that traditionally in linguistic description even outside Indo-European, when a paradigm has been identified as heteroclitite, in practice this is likely to be because it displays the combination of properties which characterizes the ancient Indo-European phenomenon – namely phonologically unaccountable stem alternation in the noun,

implicating a paradigmatic opposition between ‘direct’ and ‘oblique’ in some sense.⁵ This conception is endorsed, for example, in Lehmann and Moravcsik’s (2000) article on noun inflection in the HSK International Handbook on Morphology (Booij et al. 2000): this illustrates the notion of *heterocclisis* by means of the behaviour of Tamil *maram* ‘tree’, which builds all its case forms outside the nominative (in the singular) on an oblique stem *maratt-*. Along the same lines, Menges (1978: 374) uses the term to refer to a pattern of allomorphy in Nanai (South Tungusic) whereby some nouns make use of a special stem before overt case markers, e.g. *dèrèl* ‘upper part’, whose ‘declension stem’ is *dèrè-g-*. Note, moreover, that the linguistic example selected by Conrad 1985 in the definition above, Russian *mat* ‘mother’, is also one in which the paradigmatic split opposes a common nominative-accusative form (*mat*) to the rest of the paradigm (built on the stem *mater’-*), even though the similarity to the behaviour of the ancient *r/n*-stems is entirely coincidental (Shevelov 1964: 164) – and even though this distribution over the paradigm is not mentioned as part of the definition.

In fact, however, outside Indo-European scholarship ‘heterocclisis’ has not generally enjoyed a particularly high profile: there is no entry for it (or ‘heterocclite’) in the Encyclopedia of Language and Linguistics (Brown and Anderson 2006), the Encyclopedia of Language and Linguistics (Asher and Simpson 1994), or the International Encyclopedia of Linguistics (Frawley 2003), nor do the terms appear in the index of Katamba 2004, a six-volume collection on critical concepts in morphology. It is likely that this relative obscurity of ‘heterocclisis’ as an item of general linguistic terminology is largely due to the fact that both the general meaning, as relayed in some manuals, and the more specific ‘Indo-European’ use to

⁵ ‘In some sense’, because note that for this characterization to apply to the ancient Indo-European case itself the accusative singular has to be treated as ‘direct’ and the nominative plural ‘oblique’.

which it has sometimes been put are covered by other, more familiar concepts such as suppletion, allomorphy or simply irregularity, all of which are equally applicable in the nominal domain and outside it.

It could thus be said that ‘heterocclisis’ is a venerable technical term, long in use with reference to patterns of alternation within the paradigm identifiable as exceptional against the backdrop of the inflectional system as a whole – but one whose potential broadness, but effective restriction to the description of a specific pattern which falls under the wider umbrella of stem allomorphy and cannot easily be generalized outside the nominal domain have combined to see it languishing somewhat, at least outside Indo-European philology. At the same time, as we have said, recent years have seen an increasing scholarly interest in exceptionality and marginal behaviour in morphological systems, and in the typological treatment of such phenomena, for which a standardized technical vocabulary is required. In this context, the privileging of a different facet of the term’s classical use seen here and in the contemporary morphological literature, which allows its generalization beyond the noun and provides a way to discuss alternations in inflectional class affiliation over the paradigm, fulfils a valuable role.

Among recent work which concentrates on heterocclisis in the sense intended here, however, differences of emphasis and approach still remain which leads scholars to draw the boundaries of the phenomenon in very different places. In sections 2.4 and 2.5 below, I turn to a presentation of two major theoretical treatments of heterocclisis which emerged independently at around the same time (and largely cemented the modern association of the term strictly with inflectional class behaviour), namely Noyer 2005 and Stump 2006, which diverge in numerous ways. One striking disagreement is found in the fact that Noyer treats as examples of

heteroclisis only *unsystematic* phenomena, particular to individual (or very few) lexemes in a language, which are not to be taken as constituting a coherent group: the behaviour of Latin *domus* would be one example of heteroclisis in these terms. Noyer contrasts this with cases in which a significant number of lexemes share a pattern of inflectional behaviour of this type and thus cannot be treated as showing irregularity in quite the same way: for example, a group of masculine nouns which largely pattern with a declension mostly containing feminines, but pattern in certain cells of the paradigm with a different declension associated with masculines. For Noyer, in such cases we are witnessing not heteroclisis but ‘systematic mixed inflection’. For Stump, meanwhile, both are equally instances of heteroclisis, both meeting the definition of the phenomenon given recently by Stump and Finkel (2013: 175): ‘A heteroclite lexeme is one whose paradigm inflects partly according to the pattern of one I[nflection] C[lass] and partly according to the pattern of another’.

This significant disparity between Noyer and Stump is connected with the differences between the theoretical accounts they provide for the behaviour treated. At the same time, however, we find a situation where the range of inflectional phenomena treated by the two scholars is recognized by both to share a fundamental core, which is that exemplified by the behaviour of *domus*, in which separate inflectional classes provide material for a single paradigm in an entirely, or almost entirely, lexically idiosyncratic way; differences arise as they consider types of inflectional behaviour with a resemblance to such cases along some dimension, but lying further and further away from the core.

In allowing for this conception as a gradient phenomenon, heteroclisis is in no way different from other types of exceptionality in inflectional morphology, for which it may equally be possible to point to central, clear-cut cases while

acknowledging a penumbra of less ‘satisfactory’ examples affiliated with the conceptual ideal to a greater or lesser extent along different dimensions: that morphological description and comparison, and indeed linguistic typology more generally, can benefit from a recognition of the gradient nature of crucial descriptive notions is an important tenet of the programme of Canonical Typology to be discussed below (section 2.3 with references), and many of the works to have touched on heteroclasia in recent years have done so from a ‘canonical’ perspective. One upshot of this view is that we should not be surprised to come across morphological behaviour which is hard to categorize as representing one linguistic phenomenon rather than another. For example, Corbett (2007b: 35-38), following Comrie (2001: 381-83), points to two nouns in Tsez (East Caucasian), *xexbi* ‘child(ren)’ and *y^ʃanabi* ‘woman/women’, whose idiosyncratic inflectional behaviour could be described in terms of deponency, in the general sense: their morphosyntactically singular forms (as evidenced by agreement) inflect as regular plurals (in *-bi*), meaning that dedicated plural morphology is here being used for the ‘wrong’ feature. At the same time, as the morphosyntactically plural forms of these nouns also inflect as regular plurals (and are therefore identical to the singulars in every case), they could equally be described as showing syncretism for number. This case is thus difficult to classify as ‘really’ exemplifying syncretism as opposed to deponency or the reverse – instead falling somewhere between, but distant from, the most securely recognizable instances of the two.

The potential for such intermediate instances to exist, in the case of deponency and syncretism, is due to the fact that both of these phenomena can involve the effacement, in terms of morphological realization, of distinctions which must nonetheless be recognized at the morphosyntactic level; even so, the distinction

between deponency and syncretism remains well-motivated, and in general it is clear when it is appropriate to identify one or the other. Likewise, there are commonalities between heteroclisis and other inflectional phenomena, which nonetheless remain distinct from it even though indeterminate cases can be imagined. For example, consider heteroclisis in comparison to semi-deponency (for voice) in the Latin verb system. In Latin deponency proper, a phenomenon concerning the expression of the active/passive distinction whose prominence as a topic of linguistic theorizing means that it has lent its traditional name in Latin grammar to the cross-linguistically identifiable type of form-function mismatch it exemplifies (as in Tsez above, and cf. Baerman et al. 2007 more generally), fully deponent verbs make use of practically all the morphology proper to the passive portion of the paradigm, while retaining the syntactic characteristics of active verbs: see e.g. Flobert 1975 for a comprehensive treatment. Thus a deponent verb such as *(eum) sequor* ‘I follow (him)’, despite governing a direct object, exceptionally shows the inflection expected for a passive such as *dūcor* ‘I am led’: the same applies in the perfect, which for deponent verbs is realized by a participial periphrasis just as in the passive (*secūtus sum* ‘I followed’, cf. *ductus sum* ‘I was led’). In the much smaller group of semi-deponent verbs, meanwhile, unexpected passive morphology is found only in a section of the paradigm (specifically, the section where the perfect periphrasis is found in regular passives). Thus semi-deponent *fīdō* ‘I trust’ shows active morphology in the present and in the majority of its paradigm, but its perfect *fīsus sum* and related forms have the appearance of passives. Semi-deponency is thus a type of deponency restricted not only to particular lexemes but to certain cells in the paradigm of those lexemes as well. This state of affairs is illustrated in Table 2.1, which gives only the present and perfect indicative forms of *fīdō* alongside the corresponding forms of *dūcō*.

Table 2.1 Semi-deponent inflection of Latin *fidō* ‘trust’ (cf. Allen and Greenough 1903: §192)

	active <i>dūcō</i> ‘lead’		semi-deponent <i>fidō</i> ‘trust’		passive (<i>dūcōr</i> ‘be led’)	
	present	perfect	present	perfect	present	perfect
1sg	<i>dūcō</i>	<i>dūxī</i>	<u><i>fidō</i></u>	<i>fīsus sum</i>	<i>dūcor</i>	<i>ductus sum</i>
2sg	<i>dūcis</i>	<i>dūxistī</i>	<u><i>fidis</i></u>	<i>fīsus es</i>	<i>dūceris</i>	<i>ductus es</i>
3sg	<i>dūcit</i>	<i>dūxit</i>	<u><i>fidit</i></u>	<i>fīsus est</i>	<i>dūcitur</i>	<i>ductus est</i>
1pl	<i>dūcimus</i>	<i>dūximus</i>	<u><i>fidimus</i></u>	<i>fīstī sumus</i>	<i>dūcimur</i>	<i>ductī sumus</i>
2pl	<i>dūcitis</i>	<i>dūxistis</i>	<u><i>fiditis</i></u>	<i>fīstī estis</i>	<i>dūciminī</i>	<i>ductī estis</i>
3pl	<i>dūcunt</i>	<i>dūxērunt</i>	<u><i>fidunt</i></u>	<i>fīstī sunt</i>	<i>dūcuntur</i>	<i>ductī sunt</i>

In this table, underlined forms pattern with the active of regular verbs, and forms in bold pattern with the passive. It can be compared with Table 1.2 above, which treats the heteroclitite inflection of Czech *pramen* ‘spring’, and is deliberately laid out in such a way as to facilitate the comparison. As in the Czech case, *fidō* represents a small group of lexemes whose paradigm is split into two zones, reflecting distinct patterns of inflectional behaviour which, while unexceptional in their own right, do not usually mix: part of what distinguishes *fidō* from normal Latin verbs is its surprising juxtaposition of these two patterns, as was recognized by the Latin grammarians themselves (e.g. Priscian 2.420.9).

The similarity between the two morphological phenomena seen here is thus clear. However, there remains a principled difference between the heteroclitite of *pramen* and the semi-deponency of *fidō*. In the inflection of the latter, what is unusual is not the fact that *fidō* itself (etc.) and *fīsus sum* (etc.) are to be found in the paradigm of the same verb: this much is entirely regular, just as the paradigm of *dūcō* contains *ductus sum*. The crucial point is that the formal opposition visible between the first and third set of forms in this table, headed by *dūcō* and *dūcor* respectively, can be identified on the basis of the behaviour of Latin verb inflection as a whole as serving a morphosyntactic function, namely the distinction between active and passive voice: voice is itself a dimension of the verb paradigm, and one orthogonal to

tense. Meanwhile, the paradigm of *fidō* employs the formal opposition, but does not thereby realize the morphosyntactic distinction which can generally be seen to correspond to it. This is not the case for *pramen*, which likewise ‘compromises’ a formal opposition in some sense, but one which usually distinguishes the inflection of separate lexemes rather than different parts of the same inflectional paradigm. The difference between the two cases thus concerns the different roles which can be attributed to the patterns of morphological marking involved, on the basis of the wider inflectional systems in which they are found.

It is possible to imagine a system where the identification of a pattern of inflectional behaviour as semi-deponent or heteroclite would be less easily settled: this issue would arise in a case where it was not clear whether the formal opposition involved should indeed be taken as *purely* formal, as in the Czech case, or embodying a distinction within the paradigm, as in Latin as presented here. Here, for example, it has been taken for granted that *dūcō* ‘I lead’ and *dūcor* ‘I am led’ belong to the inflectional paradigm of a single verb, which is able to inflect for voice: the identification of voice as an inflectional feature in Latin is why *dūcor* is presented in brackets at the head of its column in the table above (it is not a separate lexeme from *dūcō*), which is the point where the parallelism between this and Table 1.2 breaks down. Meanwhile, although the combination of inflectional behaviour shown by *pramen* is just as lexically restricted as that of *fidō* and is just as clearly divided between two more normal patterns of inflectional behaviour, there is no functional opposition associated with the distinction (exemplified by *pokoj* ‘room’ vs. *most* ‘bridge’) involved. In a system where the status of the opposition involved was less clear than in the Latin or Czech examples, it would be correspondingly harder to

distinguish between heteroclisis and semi-deponency.⁶ But the potential existence of such hard intermediate cases does not alter the fact that these are fundamentally two separate phenomena, just as it is appropriate to identify deponency as a whole as essentially different from suppletion despite the Tsez case noted by Comrie and Corbett. In the Latin case specifically, as long as we are prepared to treat the relationship between active and passive as inflectional – which is why it is possible to identify voice-deponent⁷ verbs for Latin at all – there is no difficulty in recognizing semi-deponency as opposed to heteroclisis here.

The distinction between heteroclisis and deponency more generally is also made clear by the fact that the two can be observed to interact, producing paradigmatic distributions which are different from those resulting from either heteroclisis or (semi-)deponency alone. Table 2.2 illustrates this possibility with data from Gothic drawn from Baerman 2007 (following Birkmann 1987). With few exceptions, Gothic verbs can be divided into ‘strong’ and ‘weak’ classes (a characteristic feature of the Germanic verb system), exemplified here by *greipan* ‘grip’ and *haban* ‘have’ respectively: among other things, they differ in their preterite formation, which makes use of a dental suffix (-t/-d/-s) in the weak verbs and one of several patterns of stem-vowel alternation (‘ablaut’) in the strong verbs, the two also selecting different personal endings. But the fourteen verbs known as ‘preterite-presents’, exemplified here by *witan* ‘know’, fall outside this dichotomy: while their preterite forms follow the pattern seen in the weak verbs, their present tense – as their name suggests – is produced in accordance with the *preterite*

⁶ The inflectional status of voice in Latin itself has in fact been challenged, although on grounds that I find unconvincing: see section 3.1, fn. 49.

⁷ ‘Voice-deponent’, because the extension of the concept of deponency to cover features beyond voice means that it is possible to identify at least ‘tense-deponent’ verbs for Latin as well, such as *meminī* ‘I remember’, *ōdī* ‘I hate’, which inflect as perfect tense forms but have present value (Sihler 1995: 565).

formation of the *strong* verbs, with regard both to ablaut and characteristic preterite endings. They are thus deponent with regard to tense, using distinctive preterite morphology in the present; but they are also heteroclite, compromising the otherwise solid lexical opposition found between strong and weak verbs in Gothic. From one point of view, heteroclisism can be thought to have come ‘to the rescue’ of the inflection of the preterite-present verbs here: with strong preterite inflection exceptionally marking present tense value in these verbs, the use of weak inflection allows for the preterite itself to be realized unambiguously as well.

Table 2.2 Heteroclisism and deponency in Gothic preterite-presents (Baerman 2007: 17)

	strong <i>greipan</i> ‘grip’		preterite-present <i>witan</i> ‘know’		weak <i>haban</i> ‘have’	
	present	preterite	present	preterite	present	preterite
1sg	<i>greip-a</i>	<i>graip</i>	<i>wait</i>	<i>wis-sa</i>	<i>hab-a</i>	<i>habai-da</i>
2sg	<i>greip-is</i>	<i>graip-t</i>	<i>wais-t</i>	<i>wis-seis</i> ⁸	<i>haba-is</i>	<i>habai-des</i>
3sg	<i>greip-ip</i>	<i>graip</i>	<i>wait</i>	<i>wis-sa</i>	<i>haba-ip</i>	<i>habai-da</i>
1pl	<i>greip-am</i>	<i>grip-um</i>	<i>wit-um</i>	<i>wis-sedum</i>	<i>hab-am</i>	<i>habai-dedum</i>
2pl	<i>greip-ip</i>	<i>grip-up</i>	<i>wit-up</i>	<i>wis-sedup</i>	<i>haba-ip</i>	<i>habai-dedup</i>
3pl	<i>greip-and</i>	<i>grip-un</i>	<i>wit-un</i>	<i>wis-sedum</i>	<i>hab-and</i>	<i>habai-dedun</i>

Lastly, it is worth noting that while heteroclisism, in the sense relevant here, is not concerned with stem shape directly but with inflectional class affiliation, cases do of course arise where both of these issues are implicated at once in the exceptional behaviour of a given lexeme. A particularly clear example can be seen in the inflection of the Romanian verb *a fi* ‘be’. Both imperfect and perfect formations in Romanian may differ according to inflectional class: to anticipate the more detailed exposition of Romanian verb morphology provided at the start of chapter 4, it can be stated here that imperfect formations in *-am* as opposed to *-eam* (etc.) are characteristic of an inflectional class labelled the *first* conjugation, exemplified in

⁸ This form of *witan* is only attested once, so it is unclear whether the ending *-eis* is systematic; other preterite-present verbs show *-es* here as would be expected on the basis of the weak preterite forms (Birkmann 1987: 100).

Table 2.3 by *a lucra* ‘work’, while perfect formations making use of a formative in *s* are characteristic of (a subclass of) the *third* conjugation, exemplified here by *a rupe* ‘tear’. As the table shows, however, the verb *a fi* exceptionally displays a perfect of this third conjugation type while this affiliation is not shared by its imperfect forms, which must be attributed to the first. At the same time, the stems on to which these tense formations are built also show an obvious opposition according to the same split in the paradigm: heteroclisis in this case is accompanied by the strong suppletion visible between the stem alternants *er-* [jer-] and *fu-*, irreducible on phonological grounds.

Table 2.3 Heteroclisis and stem alternation in Romanian *a fi* ‘be’ (Gönczöl-Davies 2007: 96-101)

	1 st conjugation <i>a lucra</i> ‘work’		(heteroclite) <i>a fi</i> ‘be’		3 rd conjugation <i>a rupe</i> ‘tear’	
	imperfect	perfect	imperfect	perfect	imperfect	perfect
1sg	<i>lucr-ám</i>	<i>lucr-ái</i>	<i>er-ám</i>	<i>fu-s-éi</i>	<i>rup-eám</i>	<i>rup-s-éi</i>
2sg	<i>lucr-ái</i>	<i>lucr-áși</i>	<i>er-ái</i>	<i>fu-s-éși</i>	<i>rup-eái</i>	<i>rup-s-éși</i>
3sg	<i>lucr-á</i>	<i>lucr-ă</i>	<i>er-á</i>	<i>fú-s-e</i>	<i>rup-eá</i>	<i>rúp-s-e</i>
1pl	<i>lucr-ám</i>	<i>lucr-ărăm</i>	<i>er-ám</i>	<i>fú-s-erăm</i>	<i>rup-eám</i>	<i>rúp-s-erăm</i>
2pl	<i>lucr-áți</i>	<i>lucr-ărăți</i>	<i>er-áți</i>	<i>fú-s-erăți</i>	<i>rup-eáți</i>	<i>rúp-s-erăți</i>
3pl	<i>lucr-áu</i>	<i>lucr-ără</i>	<i>er-áu</i>	<i>fú-s-eră</i>	<i>rup-eáu</i>	<i>rúp-s-eră</i>

It goes almost without saying that the verb *a fi* in Romanian, as might be expected for a verb with this meaning, also shows types of exceptionality which are less ‘neat’ and less easy to categorize against the background of the wider inflectional system, and indeed its full complexity even in these parts of the paradigm is understated in Table 2.3, as it also possesses a parallel set of perfect forms (*fui*, *fuși* etc.: Gönczöl-Davies 2007: 101), meaning that the verb is *overabundant* in this tense (for this label cf. Thornton 2011). However, to identify heteroclisis here brings out the point that imperfect *eram* and perfect *fusei* are not purely irregular: they show tense marking, in this case coherent across the person-number forms, which patterns with that of substantial numbers of unexceptional Romanian verbs – and the exceptionality lies in

the fact that imperfects in *-am* and perfects in *-s-ei* can otherwise be recognized as proper to *different* lexemes given the makeup of the inflectional system as a whole, just as *er-* and *fu-* can be recognized as distinct stems in the context of Romanian (morpho)phonology.

More generally, the identification of all the types of inflectional phenomena touched on above depends on the knowledge of what constitutes the ‘normal’ practice of the wider system involved. Thus deponency can only be recognized once it has been established that there is an expected link between a particular form (such as *-or* on Latin verbs or *-bi* on Tsez nouns) and a particular function (passive or plural value): if there is no such regular link, we have no grounds for identifying deponency. Similarly, to talk of syncretism implies a prior model of the cells available in principle in the inflectional paradigm: if singular and plural are never distinguished inflectionally for a given part of speech, then the question of number syncretism does not arise. The same considerations apply in the case of heteroclisis. To say that a lexeme’s inflection is heteroclite is also to make a claim about the presence of, in some sense, ‘pre-existing’ inflectional patterns which that lexeme goes against, in the same way as semi-deponent Latin *fidō* goes against what is otherwise known about the expression of voice in Latin verb inflection.

In common with others who make use of the term *heteroclisis* in reference to the phenomenon treated here, I have discussed the context in which it can be identified as a matter of *inflectional class* affiliation. In my treatment of the Latin and Romance cases to follow, I identify heteroclisis with respect to the four inflectional classes of the verb known traditionally as the Latin *conjugations*, together with their descendants in the daughter languages (as seen just above in my discussion of Romanian, where *a rupe* and *a lucra* are taken to belong to the third

and first conjugations respectively and descend from verbs in the diachronically correspondent conjugations of Latin, namely *rumpere* and deponent *lucrārī*).

But the notion of inflectional class itself remains to be explored more thoroughly. Below, in section 2.3, I return to the relationship between heteroclis and other types of morphological exceptionality in an introduction to the Canonical Typology programme, which aims to set up an abstract typological framework in which to situate the wide range of attested inflectional phenomena: I will suggest that, although it has not yet been assigned its appropriate place in this framework, heteroclis is itself susceptible to a ‘canonical’ treatment. But I turn now to consider the phenomenon of inflectional class itself, upon which the notion of heteroclis in question here is ultimately based: in a language with none of the ‘orderliness’ in inflectional realization across the lexicon which inflectional classes are posited in order to capture, neither would there be any basis for identifying *heteroclit* lexemes, or groups of lexemes, compromising such orderliness.

2.2 Inflectional diversity and inflectional classes

I have not dwelt on the idea of inflectional classes up to this point, because ‘[o]n one level we know what an inflectional class is’ (Corbett 2009a: 1): since, in attested inflectional systems, different lexemes belonging to the same part of speech are commonly found to inflect differently in order to express the same distinctions with regard to morphosyntactic features, in ways which cannot be entirely ‘blamed’ on other parts of the grammar such as the phonology or syntax, we identify those lexemes which pattern together in this respect as sharing membership in a single more or less arbitrarily constituted group, of interest to the morphology, known as an inflectional class. This practice underlies preliminary definitions along the lines of Aronoff’s (1994: 64) ‘An inflectional class is a set of lexemes whose members each select the same set of inflectional realizations’, but also explains why little use is made of the notion of inflectional class not only where allomorphy is absent, but also where it exists and can be linked reliably with extramorphological factors such as phonological conditioning:⁹ in such cases, to posit a level of classification ‘situated’ in the morphology often adds little to our conception of what is going on. The examples adduced so far in order to illustrate the concept of heteroclasia and its relationships with other morphological phenomena have all taken for granted an appreciation of this basic role which may be played by such classes in the way inflectional behaviour applies to the lexemes of a given language.

But even if the existence of inflectional allomorphy is taken for granted as a brute fact to be reckoned with in the treatment of some languages, it should be

⁹ In Turkish, for example, the 3sg possessive marker takes the form *-I* following a consonant and *-sI* following a vowel (Lewis 2000: 37f.; *I* signals a high vowel subject to front/back and rounding harmony). Although the presence or absence of *s* is not an automatic phonological alternation, being linked to this particular morphological context, the notion of inflectional class is not generally invoked here to distinguish the two groups of lexemes involved.

acknowledged that there is no *logical* reason why the notion of inflectional class should empirically prove to be a useful one. Given the existence of various allomorphs to perform each of various inflectional tasks, every lexeme sensitive to the same set of inflectional distinctions could well simply select its own set of allomorphs over the paradigm, in such a way that every combination of allomorphs was attested up to the limit of the total number of lexemes involved (if smaller than the total number of combinations).¹⁰ In such a system, it would technically be possible to identify inflectional classes, but they would be devoid of any linguistic interest: the presence or absence of a given inflectional pattern would say nothing about it and would be unconnected to the presence or absence of any other, and every combination of lexical meaning and inflectional function would need to be learned and stored individually.

In fact, however, this state of affairs is not what we find. Instead, we tend only to find a small fraction of the patterns of inflectional behaviour which could in principle exist on the basis of the observed inflectional resources; the upshot of this is that principled presences and absences are identifiable in the combinations of allomorphy available to any part of speech in an inflecting language, which is precisely why the idea of treating lexemes as belonging to different inflectional classes is ever treated as more than simply an accounting device for listing the full complement of forms available to each lexeme individually. This empirical observation about inflection cross-linguistically can obviously be linked with questions about the potential limits of memory and the learnability of such systems, a point to which I return below in section 2.2.3.

¹⁰ 'Purely in terms of [inflectional] resources it would make no difference whether Latin had six, seven, eleven, or 1,851,776,640 declensions of nouns' (Plank 1991: 29), where the number 1,851,776,640 is calculated (on p. 27) by multiplying out all the possible combinations of the allomorphs actually attested in Latin according to one method of analysis.

But morphologists in different research traditions address this empirical observation of inflectional structure in different ways, contrasting in their approach with regard to the details of inflectional classification and privileging various means to account for the regularities and restrictions on inflectional diversity that are actually found. In the following sections (2.2.1 and 2.2.2), I look in some detail at two such research traditions, associated particularly with the work of Andrew Carstairs(-McCarthy) and Wolfgang Wurzel respectively, which are intended to exemplify two broad approaches to the question of constraints to diversity in inflectional systems. These sections, together with the subsequent discussion (2.2.3), bring out some of the detailed issues which surround the notion of inflectional class. Moreover, without focusing specifically on the concept of heteroclasia, or indeed making use of the term in the sense relevant here, the theoretical approaches presented also have things to say about the kinds of inflectional class mixture over the paradigm which this thesis addresses.

One influential variety of theoretical response to this observed characteristic of inflectional behaviour has been to propose the existence of a general cross-linguistic constraint applying to the distribution of exponents within inflectional systems, which automatically brings about a much higher degree of predictability within paradigms than could be expected in the absence of the constraint. Such approaches are associated particularly with the research tradition inaugurated by Carstairs-McCarthy in work making use of the notion of ‘paradigm economy’ (introduced in Carstairs 1983), and it is to his work that I turn first of all.

2.2.1 Carstairs-McCarthy and affix economy

As Blevins notes (2004: 47), Carstairs 1983 appears to have been the first to observe that there is no intrinsic reason why we should expect to find any level of

interdependence between the inflectional exponents found in the different cells of a lexeme's paradigm: as pointed out just above, given the existence of inflectional allomorphy in a language, it would logically be possible for each lexeme to inflect using a different combination of exponents across its cells, thereby rendering the notion of inflectional class otiose within the workings of the system. The empirical fact that nothing approaching such a state of affairs is observed in genuine spoken language is hence explained by Carstairs 1983, at least in part, in terms of what is effectively an *extrinsic* constraint, the *Paradigm Economy Principle* (henceforth PEP). To a first approximation (subject to refinements below) and in its strongest form, this states that the number of distinct 'paradigms', i.e. inflectional classes, available to lexemes belonging to a given part of speech will in fact be no more than the *minimum* logically possible given the array of exponents found. That is to say, according to this principle the number of inflectional classes will not exceed the number of different exponents available to the paradigm cell which makes use of the most exponents altogether across the lexicon: situations where 'paradigm economy' of this kind is not observed will be unstable and susceptible to morphological change, thus ensuring that the number of inflectional classes available to a system is always 'at or close to the minimum logically compatible with [its] inflectional resources' (Carstairs 1983: 127). It can be noted here that the operation of such a principle should clearly disfavour the survival of heteroclite paradigms, which by definition display no distinctive morphology of their own, though this must be qualified below.

As an illustration of the empirical effect which this principle is intended to account for, consider the behaviour of verbs in Burmeso, a language of Western New Guinea (Donohue 2001), which fall into two conjugations as exemplified by *-ihi-*

‘see’ and *-akwa-* ‘bite’ in Table 2.4 below. Burmeso verbs show cumulative number and gender agreement with the object by means of prefixes; two numbers and six genders (labelled I-VI in the table) are distinguished.

Table 2.4 Verbal inflectional classes in Burmeso (Corbett 2009a: 9; Donohue 2001: 100, 102)

	1 st conjugation <i>-ihi-</i> ‘see’		2 nd conjugation <i>-akwa-</i> ‘bite’	
	singular	plural	singular	plural
I	<i>j-ihi-</i>	<i>s-ihi-</i>	<i>b-akwa-</i>	<i>t-akwa-</i>
II	<i>g-ihi-</i>	<i>s-ihi-</i>	<i>n-akwa-</i>	<i>t-akwa-</i>
III	<i>g-ihi-</i>	<i>j-ihi-</i>	<i>n-akwa-</i>	<i>b-akwa-</i>
IV	<i>j-ihi-</i>	<i>j-ihi-</i>	<i>b-akwa-</i>	<i>b-akwa-</i>
V	<i>j-ihi-</i>	<i>g-ihi-</i>	<i>b-akwa-</i>	<i>n-akwa-</i>
VI	<i>g-ihi-</i>	<i>g-ihi-</i>	<i>n-akwa-</i>	<i>n-akwa-</i>

As is clear from this table, each gender/number value for object agreement in Burmeso can be expressed by two different prefixes, e.g. IIIsg agreement can be realized by *g-* or by *n-*. Trivially, every cell of the verb paradigm of Burmeso, taken in the abstract, has the ‘most’ rival exponents available to it, namely precisely two. If any given verb were able to select either of the available exponents in each of its cells, independently of the exponent selected by that verb in any other cell, the language could potentially display a maximum of 2^{12} (= 4096) distinct object agreement patterns.¹¹ Instead, however, the logical minimum of only two distinct inflectional classes is observed, a finding in accordance with the PEP.

While this highly streamlined system neatly embodies the notion of economy in question here, a slightly more involved example is required to bring out the nuances of the way in which the PEP is envisaged to apply more generally and the nature of the entities to which such ‘economy’ is supposed to apply. I follow Blevins (2004: 50-53) and Müller (2007: 169f.) in using the example of Russian nominal

¹¹ Even if the observed paradigmatic pattern of syncretism were taken as fixed on some independent grounds, 2^3 (= 8) patterns could be found, rather than only the two attested.

declensions for this purpose, although for expository reasons all three presentations differ from each other and from Carstairs's own (1987: 60-62) in the range of (somewhat idealized) Russian data they present.

Table 2.5 Inflectional suffixes of the Russian noun (after Brown and Hippisley 2012)

	1 st decl. 'hard' (m) <i>zakon</i> 'law'	1 st decl. 'soft' (m) <i>rubl'</i> 'rouble'	2 nd decl. <i>karta</i> 'map'	3 rd decl. <i>noč'</i> 'night'	4 th decl. (n) <i>slovo</i> 'word'
nom sg	-∅	-∅	-a	-∅	-o
acc sg	-∅	-∅	-u	-∅	-o
gen sg	-a	-a	-i	-i	-a
dat sg	-u	-u	-e	-i	-u
instr sg	-om	-om	-oj	-ju	-om
prep sg	-e	-e	-e	-i	-e
nom pl	-i	-i	-i	-i	-a
acc pl	-i	-i	-i	-i	-a
gen pl	-ov	-ej	-∅	-ej	-∅
dat pl	-am	-am	-am	-am	-am
instr pl	-ami	-ami	-ami	-ami	-ami
prep pl	-ax	-ax	-ax	-ax	-ax

Table 2.5 presents the inflectional suffixes available in the Russian nominal system, as they are distributed over the paradigms of five nouns which between them represent the major inflectional patterns found, to the exclusion of certain complexities which are not germane to this presentation¹² (and with the effects of automatic phonological processes undone; cf. Brown and Hippisley 2012: xx-xxii, 126). As is immediately clear from the table, while these five nouns all differ from one another to some extent in their inflectional behaviour, they are not all distinct in every part of the paradigm, unlike the two representative verbs seen in Burmeso above: in fact, as illustrated here, some suffixes (such as the prepositional plural *-ax*) apply to all Russian nouns. But more crucially in the present context, there is no case/number value which is marked by as many as five distinct exponents: the

¹² For example, inflectional distinctions correlated with animacy have been omitted (cf. Carstairs 1987: 60f.), as has the existence of a 'hard/soft' distinction in the neuter fourth declension in parallel with that seen in the first.

maximum number is three, as seen in the nominative, accusative, dative and instrumental singular and in the genitive plural. This inflectional system would thus appear to counterexemplify the PEP.

However, it is notable that the facts laid out in Table 2.5 can be brought into line with the PEP if one is willing to entertain an analytical practice which ‘consolidates’ some of the inflectional classes seen here into a smaller number of superordinate units – a practice which, as Carstairs points out (1987: 68f.), is commonly treated as uncontroversial in similar circumstances and is undertaken for reasons independent of the PEP. The exponents shown by nouns patterning with *zakon* ‘law’ (m.) and *rubl’* ‘rouble’ (m.) differ in only one cell of the paradigm, the genitive plural (-*ov* vs. -*ej* respectively), and over the lexicon this distinction correlates reliably with the phonology of the stem-final consonant in nouns showing these inflections: -*ej* appears when this consonant is ‘soft’, i.e. palatalized (transcribed with an apostrophe as in *rubl’*), and -*ov* appears when it is ‘hard’, i.e. unpalatalized. Accordingly, although the two endings are generally taken to be phonologically irreducible to a single element (as is assumed here, though cf. Bailyn and Nevins 2008), it is standard practice to treat *zakon* and *rubl’* as exemplifying ‘hard’ and ‘soft’ variants respectively of a single ‘first declension’, as indeed is shown by the traditional labels used in the table above.

This seems to capture a salient generalization concerning Russian nominal inflection, and indeed it is seen as so innocuous to factor out the morphological effects of palatalization that they are disregarded by Blevins (explicitly) and Müller (implicitly) even in their discussion of the very question at issue here, namely the number of ‘distinct paradigms’ relevant to the PEP in Russian nouns. However, by the same general process of abstracting away from distinctions for which

independently established factors can be held responsible, it is also possible to consolidate this unified first declension with what is labelled as the fourth declension in Table 2.5. As is signalled there, lexemes patterning with *zakon* and *rubl'* are masculine, and lexemes patterning with *slovo* are neuter: the inflectional differences between them are thus predictable on the basis of gender. If these three columns of the table are all consolidated into one, it emerges that the PEP is now satisfied: the three-way partition of the nominal lexicon which remains is 'licensed', for example, by the existence in the instrumental singular of the three suffixes *-oj*, *-ju* and *-om* (characterizing the second declension, the third declension, and the newly combined first and fourth declensions respectively).

In the terms of Carstairs 1983 and subsequent work in the same tradition, the paradigms of *zakon*, *rubl'* and *slovo* together constitute a 'macroparadigm'; within a macroparadigm, any inflectional differences 'either can be accounted for phonologically, or else correlate consistently with differences in semantic or lexically determined syntactic properties' such as gender (Carstairs 1987: 69), and it is to such *macroparadigms*, the units of inflectional organization that remain once the consolidation process has been completed, that the PEP should be taken to apply.

While the establishment of macroparadigms raises questions about the theoretical status of the 'macroinflections' that they entail (which might be expressed in the form e.g. $\{-i, -a\}$, following Blevins (2004: 52), for the nominative and accusative plural of the first/fourth declension macroparadigm above), in itself it may not appear particularly tendentious to focus in this way only on what is 'irreducibly' different about the inflectional behaviour of different lexemes. We have already seen that the standard treatment of both *zakon* and *rubl'* as 'first declension' nouns already depends on this strategy, and in fact, many descriptive treatments of Russian

nouns (e.g. Gabka 1975, Garde 1980) do collapse neuter *slovo* together with these into a single declension; meanwhile, discussing the relevance of the PEP to the Latin inflectional system, Nyman (1988: 501) observes that the notion of the macroparadigm ‘conceptualizes the age-old practice... of lumping the noun types *lupus* ‘wolf’ and *bellum* ‘war’ into (a single macroparadigm called) the 2nd declension’, i.e. of unifying certain Latin masculines and neuters in a precisely analogous way.

However, the importance of the macroparadigm to the purported workings of the PEP does detract from its empirical force, especially as Carstairs-McCarthy’s deployment of the concept in individual cases can be somewhat looser than his definition above might suggest (see further below). At the same time, certain further characteristics of the full theoretical ‘package’ associated with the PEP also moderate its scope significantly. For one thing, the qualification labelled the *Slab Codicil* (Carstairs 1987: 81) licenses the existence of inflectional behaviour which does fall foul of the PEP as initially stated, just so long as the Principle is still observed within the domain of each subparadigm (‘slab’) as marked out by some morphosyntactic property ‘fundamental’ for the given part of speech, such as aspect in verbs or number in nouns. Thus, for example, in an extended discussion of German nominal inflection, in which complex ‘weak’ and ‘strong’ macroparadigms are identified which abstract over distinctions involving phonology, gender and animacy, Carstairs (1987: 241-50) raises the possibility that certain nouns (e.g. *Dorn* ‘thorn’) should best be treated as drawing their inflection from the Strong macroparadigm throughout the singular and the Weak macroparadigm throughout

the plural,¹³ thus showing a type of heteroclisis which does not contravene the PEP in the light of the Slab Codicil.

Table 2.6 Inflectional behaviour of German *Dorn* ‘thorn’ alongside representatives of ‘strong’ and ‘weak’ macroparadigms (after Carstairs 1987: 241-50)

	‘strong’		‘weak’
	<i>Tag</i> (m) ‘day’	<i>Dorn</i> (m) ‘thorn’	<i>Diamant</i> (m) ‘diamond’
nom sg	-∅	-∅	-∅
acc sg	-∅	-∅	-en
gen sg	-s	-s	-en
dat sg	-∅	-∅	-en
nom pl	-e	-en	-en
acc pl	-e	-en	-en
gen pl	-e	-en	-en
dat pl	-en	-en	-en

Just as crucially, however, the notion of inflectional exponence to which the PEP (together with its theoretical descendants) is relevant is itself restricted. In the research tradition which investigates inflection in economy terms, affixal morphology alone is important for the determination of inflectional ‘distinctness’ and the enumeration of the exponents available to the system: for example, German *Tag* ‘day’, pl. *Tage* and *Gast* ‘guest’ pl. *Gäste* are assigned to the same paradigm in Carstairs 1987 although they differ with regard to the presence of umlaut in the plural forms. As is pointed out by Blevins (2004: 55), this means that much of the inflectional predictability observed in systems chiefly articulated by non-affixal realization cannot be taken to emerge from considerations of ‘economy’ in this sense at all.

This focus on affixal exponence is inherited – along with counterparts of the macroparadigm and the Slab Codicil – by the theoretical descendant of the PEP known as the *No Blur Principle* (NBP: Carstairs-McCarthy 1994), recently re-established on new foundations as *Vocabular Clarity* (VC: Carstairs-McCarthy

¹³ For discussion of the merits of this analysis in the given case see Wurzel (1989: 222-25) alongside the original treatment in Carstairs (1987: 241-50).

2010). Although motivated somewhat differently, these proposals can be considered together in terms of their purported effects (cf. Enger 2013), which are easiest to state in terms of the original NBP. This treats ‘inflectional class membership as part of the meaning of an inflectional affix’ (Carstairs-McCarthy 1994: 741), where inflectional classes in this formulation can be identified with the macroparadigms above; it holds that, in the inflection of a given part of speech in a particular paradigm cell, every affix should either be associated with only a single class or else represent the single class-neutral form found as a default. The overall effect is to rule out ‘blur’, the association of an affix with more than one specific class. An example of a pattern outlawed by the NBP is given in Table 2.7, where letters *a-d* stand for distinct affixes and all patterns are legitimate except Pattern 5: in this pattern, neither *a* nor *b* can be identified as the default form, but neither is associated with a single class either. Blur is ruled out on the grounds that it would require the learning of a disjunctive meaning, thus going against general biases in language acquisition (Carstairs-McCarthy 1994, 1998, citing Clark 1993).

Table 2.7 A pattern outlawed by the NBP (Enger 2007: 283)

Patterns	Class 1	Class 2	Class 3	Class 4
1	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>
2	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
3	<i>a</i>	<i>a</i>	<i>a</i>	<i>b</i>
4	<i>a</i>	<i>b</i>	<i>b</i>	<i>b</i>
5	<i>a</i>	<i>a</i>	<i>b</i>	<i>b</i>

The empirical success of the PEP is taken by Carstairs-McCarthy (1994: 757) as a by-product of the validity of the more fundamental NBP, although the two principles do not sanction precisely the same patterns of affixal distribution over classes: for example, the NBP allows for the existence of a class which uses *only* default affixes, which would violate the PEP by contributing to the total number of classes without contributing to the total number of inflectional exponents. But like the PEP itself (cf.

Wurzel 1989: 210), the later NBP/VC approach is intended to point to a universally relevant *tendency* in language rather than a strict constraint; furthermore, the continued relevance of the macroparadigm and slab concepts serve to ‘defuse’ many potential counterexamples.

Nonetheless, the work of Carstairs-McCarthy and others working in the same tradition has aimed to link attested instances of morphological change, or even the absence of expected change, to the operation of these principles. Thus Enger 2007 observes, in a study of dialectal Norwegian, that while many dialects have undergone changes to the distribution of inflectional endings over the conjugations found in the verb system, the endings most likely to spread are not (as might have been expected) those originally associated with the most populous conjugation, such as present *-ar*, a development which would have fallen foul of the NBP; instead, those endings spread which already characterized more than one (less populous) conjugation, such as present *-er*, and could thus be taken as the default and be adopted by new conjugations without giving rise to ‘blur’. It must be noted, however, that the identification of present *-er* as the ‘default’ depends on Enger’s decision to treat *kjøpe* ‘buy’ (*kjøp-er*, past *kjøp-te*) and *bygge* ‘build’ (*bygg-er*, past *byg-de*) as representing two separate conjugations, although the standard view treats them as a unified conjugation whose detailed behaviour is largely predictable on phonological grounds, as Enger observes (2007: 290f.), i.e. precisely a single macroparadigm in Carstairs-McCarthy’s terms.

An illustration of the relevance proposed for economy principles, but also of the crucial nature of the macroparadigm concept for their application in concrete instances, can be provided by Carstairs’ (1984, 1988a) earlier treatment of the Latin nominal third declension in terms of the PEP, a case ‘closer to home’ in the context

of the present thesis. As he notes, the history of Latin nominal inflection saw the ‘gradual merger of Indo-European consonant-stem and *i*-stem types’ (1984: 117): the traditional identification in Latin grammar of a single, though somewhat heterogeneous entity labelled the ‘third declension’ principally aims to capture the inflectional commonalities arising from this merger, which reflect both morphological change and regular sound change (see e.g. Nyman 1987: 259-61). However, inflectional distinctions reflecting the split between these two chief lexical sources of the third declension survive through to the historical period in some parts of the paradigm, as can be seen in Table 2.8 (based on Tables I and V in Carstairs 1984: 119, 128); this presents a somewhat idealized picture of third declension behaviour in the second century BC, which abstracts away from the divergent inflection of neuter nouns and more peripheral types also generally subsumed under the third declension label. The differences between the paradigms exemplified by *rēx* ‘king’ and *ignis* ‘fire’ reflect their historical origins as ‘consonant-stem’ and ‘*i*-stem’ nouns respectively;¹⁴ in view of these differences, Risch 1977, which deals more comprehensively with the Latin declensional system as a whole, treats these two nouns as representatives of the separate declensions 3A (*rēx*) and 3B (*ignis*).

However, what is of interest in light of the PEP is the existence of a third group of nouns, exemplified here by *sors* ‘fate’, which do not pattern exclusively with either *rēx* or *ignis*. As Table 2.8 makes clear, the paradigm of *sors* (similarly *mēns* ‘mind’, *mōns* ‘mountain’ etc.) at this historical stage behaves like that of *ignis* except with regard to the nominative singular *sors* itself (for phonologically impossible **sort-s*), which does not take the form **sortis* as might have been expected on the basis of the rest of the paradigm. The historical background to this

¹⁴ Cf. Sihler (1995: 315-19).

phenomenon is known: as demonstrated by both comparative evidence and sparse Latin attestations of the earlier state of affairs, the behaviour shown by *sors* originally resulted from a process of syncope which affected *i*-stem nouns of certain phonological shapes in the nominative singular, the earlier forms being *sortis*, *mentis* etc. (Neue and Wagener 1902: 232).¹⁵ But as Carstairs 1984 points out, the synchronic result of this syncope is the existence of a ‘paradigm’ intermediate between 3A and 3B with no unique inflectional exponent of its own: there is no row in Table 2.8 in which the suffixes shown by all three columns are distinct. That is, the situation represented in Table 2.8 falls foul of the PEP.

Table 2.8 A breach of the Paradigm Economy Principle in Latin nouns (after Carstairs 1984)

	3A		3B
	‘king’	‘fate’	‘fire’
nom/voc sg	<i>rēx</i> [rēg-s]	<i>sors</i> [sort-s]	<i>ign-is</i>
acc sg	<i>rēg-em</i>	<i>sort-em</i>	<i>ign-em</i>
gen sg	<i>rēg-is</i>	<i>sort-is</i>	<i>ign-is</i>
dat sg	<i>rēg-ī</i>	<i>sort-ī</i>	<i>ign-ī</i>
abl sg	<i>rēg-e</i>	<i>sort-ī</i>	<i>ign-ī</i>
nom/voc pl	<i>rēg-ēs</i>	<i>sort-ēs</i>	<i>ign-ēs</i>
acc pl	<i>rēg-ēs</i>	<i>sort-īs</i>	<i>ign-īs</i>
gen pl	<i>rēg-um</i>	<i>sort-ium</i>	<i>ign-ium</i>
dat pl	<i>rēg-ibus</i>	<i>sort-ibus</i>	<i>ign-ibus</i>
abl pl	<i>rēg-ibus</i>	<i>sort-ibus</i>	<i>ign-ibus</i>

Carstairs takes this feature of the inflectional system as it stands in the post-syncope period to explain some of the further changes to the third declension which are observed subsequently in the history of Latin. In particular, notice that in the singular, as presented here, the ‘offending’ paradigm of *sors* only diverges from the 3A paradigm in its ablative form *sortī*. However, in the course of the second and first

¹⁵ The question whether this syncope should itself be treated as a phonological or a morphological phenomenon is a matter for debate, addressed by Carstairs (1984: 121f.; 1988a: 495f.) and Nyman (1987: 261-63) in the context of the current discussion on the relevance of the PEP, and by others (e.g. Risch 1977: 242). The fact that only the nominative singular *sortis* was affected (and not the genitive singular which itself emerges in classical Latin as *sortis*) does not necessarily point to a morphological development targeting the nominative, as the genitive ending is likely to have been *-es* rather than *-is* at the time this syncope took place (Sihler 1995: 316f.).

centuries BC, the nouns in question show a morphological shift in precisely this inflected form, giving up inherited $-ī$ in favour of an innovative $-e$ which was initially proper to 3A (cf. *rēge*): thus Lucretius, writing in the second quarter of the first century BC, uses innovative forms such as *sorte* rather than *sortī* in 172 of 179 instances of the ablative singular of relevant nouns (Ernout 1918, cited in Carstairs 1984: 122f.). This development serves to eliminate the distinction between *sors* and *rēx* throughout the singular, meaning that the PEP is no longer violated at least in that part of the paradigm. For Carstairs, this outcome is no accident: his account treats the spread of $-e$ to these new nouns as a ‘therapeutic’ move (Nyman 1987: 258) on the part of speakers, which helps to remedy the unstable situation caused by the violation of the PEP after syncope. As he acknowledges, ablative singular $-e$ also in fact spreads beyond *sors* and the like into 3B nouns ‘proper’ in the first century BC, gradually supplanting $-ī$ there too; however, the fact that these unsyncoated nouns are affected later and less comprehensively is treated as evidence for the claim that only the *sors* development in particular was motivated by economy considerations (Carstairs 1984: 122f.; 1987: 498).

Carstairs takes developments in the plural in the same spirit, in effect making use of the macroparadigm concept to suggest that here too we can observe effects which bring third declension inflectional morphology back in line with the PEP. In the genitive plural, he proposes (1984: 130) that the inherited lexical distribution of $-um$ and $-ium$, as presented in Table 2.8, gave way in this period to a situation where the ending selected by any lexeme in the third declension was largely predictable on the grounds of its phonology, or the combination of its phonology and gender, albeit according to somewhat intricate conditions. To the extent that this is true, the two

endings were no longer in competition, instead representing a single ‘macroinflection’ shared by the third declension as a whole.

However, it is in developments connected with the *accusative* plural, which also remain to be explained, that we can perhaps see the use of such ‘macroinflections’ stretched to its limits. In the accusative plural, Carstairs notes that we might expect to find the inherited 3A ending *-ēs* spread along with the ablative singular *-e*. However, given that the two changes do not in fact take place in parallel, with certain lexemes holding out against the spread of *-ēs*, he allows for those which do retain the conservative form to be treated as being individually *lexically marked* for the exceptional accusative plural *-īs* (Carstairs 1984: 133f.), apparently without thereby falling outside the macroinflection he wishes to set up. This is clearly dangerous for a principled view of what a macroinflection can be. And yet, if anything, the notion of the macroinflection presented when Carstairs-McCarthy addresses the same part of the Latin nominal system in more recent works (notably Carstairs-McCarthy 2010: 120-29) is even less restrictive, now encompassing gender, phonology and even social variation among other things. This tends to strip notions of economy of their empirical force.

2.2.2 *Inflectional diversity in Natural Morphology*

While work on economy principles in the tradition of Carstairs-McCarthy represents one prominent response¹⁶ to the observation that inflectional allomorphy does not run wild in language and can be usefully treated in terms of inflectional classes, a different and to some extent complementary approach is provided by the Natural Morphology (NM) framework, first proposed in Mayerthaler 1981 and then extended

¹⁶ Note that while I present these two approaches as distinct in order to bring out their important differences in focus, they are by no means implacably opposed: on the contrary, ‘compromise’ models have been suggested, notably by Carstairs-McCarthy himself (1991).

more substantially to inflectional morphology especially by Wurzel (e.g. 1984) and Dressler (e.g. 1985a). Including the whole of morphological behaviour within its purview, work in this tradition diverges in emphasis from the more narrowly focused work of Carstairs-McCarthy with regard to the predictability of inflectional allomorphy over the lexicon and the notion of inflectional class in general.

The fundamental sense of ‘naturalness’ with which NM is concerned follows from that employed earlier in the theory of Natural Phonology, introduced in Stampe 1969 and more explicitly in Donegan and Stampe 1979. More natural linguistic phenomena in the relevant sense are more readily acquired, more widespread cross-linguistically, and are more resistant to change than more unnatural phenomena, for reasons which in the phonological domain can be attributed to inherent characteristics of the articulatory and perceptual systems: in this sense, for example, front unrounded vowels are observed to be more natural than front rounded vowels (Wurzel 1984: 202). In the same spirit, Mayerthaler proposes (e.g. 1987: 48-50) that it is possible to identify characteristics which, all else being equal, mark out maximally natural phenomena in morphological encoding, owing to inherent features of human perception and cognition. Thus *constructional iconicity* sees a given asymmetry in semantic markedness, e.g. between positive and comparative degree in adjectives or between singular and plural in nouns, matched by a corresponding formal asymmetry such that ‘what is semantically ‘more’ is formally symbolized by ‘more’ (Wurzel 1984: 203); meanwhile, *uniformity* and *transparency* of encoding together rule out allomorphy and polyfunctionality in favour of a situation where form and function mutually imply each other within paradigms and across the lexicon (together they constitute Dressler’s (1985b: 52) ‘semiotic principle of biuniqueness’, and Wurzel’s (1984: 178) equivalent ‘Funktion-Form-Prinzip’). For

Mayerthaler, the most natural forms use iconic, uniform and transparent morphological marking, satisfying the human cognitive preference for the iconic and facilitating the task of the perceptual system (Mayerthaler 1987: 49).

It goes without saying, however, that the morphological behaviour genuinely observed in human languages always deviates, to a greater or lesser extent, from ideal naturalness as identified in these terms, even in languages characterized predominantly by the agglutinative (or isolating) morphology which these principles might be expected to favour. This can be explained at least in part by the fact that what is most natural in morphological terms may not coincide with what is most natural as regards the phonology, leading to ‘naturalness conflicts’ visible in diachrony in which neither of these components can retain the upper hand over the long term (Wurzel 1984: 29-34, 204). Dressler further points out (1985b: 53) that there exist tensions within the morphological component as well, in that agglutinative word-forms are less than perfectly natural both on the grounds of sheer length (given a proposed optimal word size of two to three syllables, gleaned from Dressler’s observations) and on semiotic grounds, because they lose out in indexicality by imposing a greater separation between a given inflection and the lexical stem it applies to.

The coexistence of multiple rival alternatives for the expression of a given morphosyntactic value, which obviously counts as a strike against semiotic biuniqueness, arises from the point of view of Natural Morphology as the upshot of such conflicts between different naturalness principles. But as Dressler observes, these conflicts have in fact given rise to the existence of (approximations to) various ideal inflectional types, each of which “sacrifices”, as it were, the naturalness of some parameters for the sake of greater naturalness in other parameters’ (Dressler

1985a: 293): that is, the inflectional types distinguished by Skalička 1979, such as ‘agglutinative’, ‘isolating’ and ‘flective’, are taken on this view to consist of relatively stable combinations of choices made along different naturalness scales. Thus below the level of *universal* naturalness investigated by Mayerthaler it is possible to posit another, *typological* level, at which there exist naturalness principles particular to these individual language types, to the extent that real languages approach them. And analogously, at a third level Wurzel recognizes principles of *system-dependent* morphological naturalness: the inflectional system applying to a given part of speech in a language will lend itself to certain generalizations (its ‘system-defining structural properties’, Wurzel 1984: 81-89, 206) which capture what is normal in that system, and the naturalness of a morphological phenomenon taken at this level is a matter of its congruity with these structural properties. The existence of inflectional allomorphy, then, while unnatural in terms of Mayerthaler’s overarching and *system-independent* principles, can be identified as natural at the typological level – in the context of certain language types – and at the level of individual inflectional systems: it characterizes, for example, both the nominal systems of Russian and Latin and more generally the *flective* type to which these and most Indo-European languages are broadly taken to belong.

But although it may be ‘natural’ in these terms, the existence of unpredictable variation across the inflection of different words belonging to a given part of speech nonetheless poses a challenge for speakers, in that it requires the memorization of each one’s morphological behaviour along with the pairing of its shape and meaning. Accordingly, Wurzel proposes (1984: 117) that wherever possible, speakers connect a word’s inflectional behaviour with some *extramorphological* characteristic of the basic word form which is taken to be stored in the lexicon anyway, i.e. some

phonological or semantic(/-syntactic) property (including such abstract features as transitivity or gender), or with a combination of such characteristics. For example, to generate the inflectional behaviour observed for Russian *karta* ‘map’ in Table 2.5 above, Wurzel would attribute to the inflectional system of Russian nouns an implicational relationship such as the following, which links the nominative singular directly with all other inflectional forms of the noun (cf. Wurzel 1984: 119):

(2.1) *karta*: /a/ in basic form (nom sg) \supset
 {/u/ in acc sg, /i/ in gen sg, /e/ in dat sg, ... /ax/ in prep pl}

Meanwhile, and only in those cases where a word’s inflectional behaviour cannot be determined on the basis of any combination of its extramorphological properties, the ‘left-hand side’ of such implications may also feature further reference forms (*Kennformen*) drawn from elsewhere in the paradigm beyond the basic lexical entry, whose inflectional characteristics allow the remaining forms to be derived: these have much in common with the principal parts of traditional language pedagogy.¹⁷ Compare (2.2), which captures the inflection of Old English *stān* ‘stone’ by means of the reference form gen sg *stānes* (Bertacca 2009: 105):

(2.2) *stān*: /Ø/ in basic form (nom sg), /es/ in gen sg \supset {/e/ in dat sg, ...}

The presence in the grammar of such implicative relationships (called paradigm structure conditions, *Paradigmenstrukturbedingungen*: Wurzel 1984: 208, 122) underpinning the inflectional behaviour of individual lexemes, the fact that they are primarily geared towards extramorphological properties, and the observation that, as a matter of system-congruity, some cells of the paradigm are more likely than others

¹⁷ To be more precise, in the categorization proposed by Stump and Finkel (2013: 29-35) traditional pedagogy most commonly makes use of *static* principal parts (representing the same cells of the paradigm for every lexeme belonging to the same part of speech), while reference forms in the Natural Morphology tradition can be identified as *adaptive* principal parts (which may differ from lexeme to lexeme, as in the Latin and German cases presented below). Note also that Bertacca’s formulation here does not in fact lay out the reference form in full but gives only its morphologically salient element, here the ending *-es*: this practice follows Wurzel’s own, as seen in the Latin examples just below.

to provide reference forms where these are necessary (e.g. the nominative plural in the German system of nominal inflection, cf. Wurzel 1990: 207) together make up Natural Morphology's account for the fact that the existence of allomorphy does not tend to give rise to limitless variation between the morphological behaviour of different lexemes. This offers a different perspective on the relative orderliness of inflection from that supplied by Carstairs-McCarthy's more 'top-down' proposal of economy constraints operating at the level of the lexicon across an entire part of speech.

Wurzel (1984: 119-21) in fact illustrates the application of paradigm structure conditions, in more complex circumstances than those just adduced from Russian and Old English, with the help of the same fragment of the Latin nominal system which was treated above in terms of its relevance to the No Blur Principle. On the basis of data drawn from Bailey (1980: 40), Wurzel discusses a historical stage (after the start of the inflectional developments provoked by the *sortis* > *sors* syncope) at which some originally '3B' nouns in nom sg *-is* have taken on inflectional behaviour characterizing '3A', as presented below in Table 2.9, which is based on Wurzel's (1984: 120) Table 2:

Table 2.9 A partial paradigm of some Latin 'third declension' nouns (after Wurzel 1984: 120)

	'3A'		'3B'	
	'king'	'citizen'	'ear'	'fire'
nom sg	<i>rēx</i>	<i>civis</i>	<i>auris</i>	<i>ignis</i>
acc sg	<i>rēgem</i>	<i>civem</i>	<i>aurem</i>	<i>ignem</i>
abl sg	<i>rēge</i>	<i>cive</i>	<i>aure</i>	<i>ignī</i>
acc pl	<i>rēgēs</i>	<i>civēs</i>	<i>aurīs</i>	<i>ignīs</i>
gen pl	<i>rēgum</i>	<i>civium</i>	<i>aurium</i>	<i>ignium</i>

This table, which omits some parts of the paradigm which are not relevant to the inflectional distinctions under discussion, is intended to lay out the different ways in

which Latin nouns juxtaposed the morphological markers available at the relevant period – the crucial point being that only certain patterns are presented as being available, while others are apparently ruled out. Thus *civis* shows a ‘3A’ ablative singular *cive* alongside a ‘3B’ genitive plural *civium*, but no word shows the opposite behaviour, in which a ‘3B’ ablative singular would accompany a ‘3A’ genitive plural. Wurzel observes that these asymmetries mean it is possible to capture the above by means of implicational *chains*, which operate only in one direction: thus the existence of a genitive plural in *-um* implies the existence of an accusative plural in *-ēs* (while the reverse is not the case), and in turn an accusative plural in *-ēs* implies an ablative singular in *-e* (but not vice versa). Meanwhile, an ablative singular in *-ī* implies an accusative plural in *-īs*, which itself implies a genitive plural in *-ium*. As a result, the behaviour seen here can be represented as being articulated by the following ‘multi-stage’ paradigm structure conditions (PSCs) linking inflectional properties with one another:

- (2.3) /um/ in gen pl \supset /ēs/ in acc pl \supset /e/ in abl sg
 (2.4) /ī/ in abl sg \supset /īs/ in acc pl \supset /ium/ in gen pl

What is more, there also exist a small number of nouns which inflect like *ignis* with the exception that they generally have an accusative singular in *-im* (Wurzel cites *puppis* ‘stern of a ship’); although the diachronic explanation for these forms is unclear,¹⁸ in synchronic terms this means that *both* of the implicational PSCs above can be extended, as an accusative singular *-im* can now be taken to imply ablative singular *-ī* (and everything that follows) but not vice versa, while ablative singular *-e* further implies accusative singular *-em*.

According to this approach, thanks to the existence of these complex PSCs in the grammar, speakers can determine the full inflectional behaviour of one of the

¹⁸ Cf. Sihler (1995: 65, 316).

relevant nouns on the basis of just one or at most two reference forms which must be stored in the lexicon alongside the basic form.¹⁹ For example, the presence of a reference form gen pl *rēgum* in the lexicon alongside *rēx* is sufficient to generate all the other forms on the basis of PSC (2.3) above; meanwhile, the lexical entry of *auris* stores the reference form acc pl *aurīs*, which implies the gen pl *aurium*, but it must also contain abl sg *aure* as a reference form (as this is not implied by any other inflected form of the noun), which itself entails that the acc sg is in *-em* rather than *-im*.

Certain details of Wurzel's presentation, which is intended to illustrate the potential complexity of such implicational relationships within the paradigm, are in fact debatable: Dressler (1985b: 56, fn. 8) disputes some of Bailey's claims as followed by Wurzel, pointing out in particular that in Caesar, for example, *ignis* has only abl sg *ignī* and acc pl *ignēs* (similarly *navis* 'ship' and others), reliably displaying a pattern of inflectional behaviour which cannot be accommodated on the basis of the PSCs above. However, he does not dispute the existence of implications of this kind within Latin noun paradigms in general, instead endorsing those presented at greater length by Risch 1977, who likewise recognizes that certain forms in the paradigm can be treated as having a special predictive or structural role.

A more thoroughly elaborated application of this approach is presented in Bittner 1985, 1996 and taken up with certain modifications by Dammell (2011: 196-210), who treats the organization of verb inflection in German in terms of PSCs designed to capture some of its overarching distributional regularities – specifically those related to the inherited Germanic distinction between 'weak' and 'strong' conjugational morphology already mentioned for Gothic above (Table 2.2), which is

¹⁹ It is not entirely clear from Wurzel's presentation what PSC or PSCs he posits to produce the remaining inflectional behaviour, which is held in common by all the nouns involved here, but the exact details envisaged are not crucial for our purposes.

reflected at several points in the paradigm. Central to Bittner's conception is the fact that 'weak' inflectional behaviour can be identified as the normal case for German verbs (i.e. *natural* in system-dependent terms) on numerous grounds: not only does it show a vastly higher type frequency over the lexicon, but it is favoured by language change, is generally applied to neologisms and nonsense forms, and is more readily acquired by children and retained by aphasic patients (Bittner 1996: 64-70; 201, cf. Wurzel 1984: 71-73). Accordingly, his model of German conjugation takes weak forms to appear by default, i.e. in cases where the verb's inflectional behaviour depends solely on its extramorphological property of *being a verb*, with no further qualifications. Meanwhile, the various morphological characteristics identified as belonging to 'strong' inflection, which are exemplified by the behaviour of *helfen* 'help' laid out in Table 2.10 below, as opposed to weak *decken* 'cover',²⁰ appear as the result of lexical specification. (In this table, the 1sg forms given should be taken to illustrate distinctions which run throughout the preterite and past subjunctive, the remaining forms differing from the 1sg in terms of their personal endings only; meanwhile, in other parts of the synthetic paradigm not provided here, such as the remainder of the present, there is no strong/weak distinction. *Ablaut* and *umlaut* are characteristic German types of vowel alternation pattern with morphological significance, also identifiable outside the verb system.)

²⁰ This presentation, with the associated PSC given below, draws on Dammel's (2011: 196f.) restatement of Bittner's generalizations, which disentangles the phenomena in question here from others which are not central to the strong/weak distinction and which, she argues, are better treated separately even on Bittner's own terms (and whose inclusion should predict diachronic developments which are not in fact attested). Dammel also (p. 198) conflates the two stages of Bittner's proposed PSC which deal with the preterite and past subjunctive, pointing out that they cannot truly be treated as standing at different points in an implicational chain. These modifications simplify the presentation of Bittner's work here but do not alter its value as an illustration of the Natural Morphology approach to inflectional diversity.

Table 2.10 Distinctive characteristics of strong inflection in Standard German (after Bittner 1996: 55-61, Dammel 2011: 196-98)

	<i>decken</i> 'cover'	<i>helfen</i> 'help'	distinctive characteristic(s) of strong inflection
	(weak)	(strong)	
2sg imperative	<i>deck(-e)!</i>	<i>hilf-Ø!</i>	<i>e/i</i> alternation; zero ending <i>only</i>
2sg, 3sg present	<i>deck-st,</i> <i>deck-t</i>	<i>hilf-st,</i> <i>hilf-t</i>	<i>e/i</i> alternation or <i>umlaut</i> with respect to basic form
preterite (1sg given here)	<i>deck-te</i>	<i>half-Ø</i>	<i>ablaut</i> with respect to basic form; lack of dental suffix; zero ending in 1sg and 3sg
past subjunctive (1sg given here)	<i>deck-te</i>	<i>hähf-e</i>	lack of dental suffix; <i>umlaut</i> with respect to preterite
past participle	<i>ge-deck-t</i>	<i>ge-holf-en</i>	<i>ablaut</i> with respect to basic form; <i>-en</i> suffix rather than dental suffix

However, as Bittner observes, German verbs do not show a neat dichotomy between those that behave like *decken* and those that behave like *helfen*: verbs are also found which diverge from default 'weak' behaviour in certain parts of the paradigm only, meaning that strong inflection in one form does not entail its presence throughout. For example, *heben* 'lift' does not pattern with *helfen* in its (relevant) imperative and present forms (*heb(e)*; *hebst, hebt*), as it displays no *e/i* alternation and does not show a 'mandatory' zero ending in the 2sg imperative; however, its preterite, past subjunctive and past participle (*hob*; *höbe*; *gehoben*) show all the 'strong' characteristics noted in Table 2.10 for these parts of the paradigm (Bittner 1996: 92). Meanwhile, the paradigm of *melken* 'milk' is weak except for its past participle *gemolken* (p. 102), which is characterized by both *ablaut* and the distinctive strong suffix *-en*.

On the other hand, the situation is not one of total inflectional 'chaos' across those verbs which do not straightforwardly follow the default, because many thinkable inflectional patterns are absent across the lexicon: for example, while a number of verbs are strong only in their past participle (like *melken*), there is no verb

which diverges from the pattern of *decken* only by displaying *e/i* alternation and an obligatory zero ending in the 2sg imperative, with no other strong characteristics elsewhere in its paradigm. In fact, it emerges that such behaviour in the 2sg imperative implies the full range of strong behaviour seen for *helfen* above. The upshot of Bittner's work is that the different patterns observed can all be captured in terms of an implicational PSC similar to those applied above to Latin nouns, which appears as follows (after Dammel's (2011: 197) formulation):

- (2.5) 2sg imperative: *e/i* alternation, zero ending
 ⊃
 2sg, 3sg present: *e/i* alternation or umlaut
 ⊃
 {preterite: ablaut, lack of dental suffix, zero ending in 1sg+3sg; AND
 past subj: lack of dental suffix, umlaut with respect to preterite}
 ⊃
 past participle: ablaut, *-en* suffix

Accordingly, the lexical entry of a verb which shows deviations from the default weak inflectional pattern need not be specified separately for each one of its strong forms, but only for that one which stands 'earliest' in the PSC given in (2.5): this lone reference form then implies the inflectional characteristics coming later in the chain, and these override the default weak forms that would otherwise have appeared (this model thus makes use of the widely recognized principle of rule ordering whereby the more specific rule trumps the more general, cf. Zwicky 1985; Bittner 1996: 181). For example, given the PSC in (2.5), the pattern of behaviour shown by *helfen* can be generated by attributing to its lexical entry the reference form 2sg imperative *hilf* (or equivalently, in a format closer to that given by Bittner, 2sg imperative: *e/i* alternation, zero ending), which generates the remaining forms. Meanwhile, the lexical entry of *melken* can be taken to specify past participle *gemolken* as a reference form, which implies nothing further according to (2.5): the remainder of the paradigm thus inflects according to the default pattern. On the other

hand, the strong past participle of *heben* is not stored as a reference form, but follows automatically from the lexical listing of a reference form belonging to the preterite or past subjunctive, such as 1sg preterite *hob* (for all these lexical representations, cf. Bittner 1996: 185).

This model of German verb inflection neatly addresses the observation that synchronically, given the existence of the various inflectional characteristics in binary opposition labelled as distinctively ‘weak’ and ‘strong’ in Table 2.10, their distribution over the paradigm of any particular lexeme may be by no means entirely predictable (as it is in the binary oppositions seen in Burmeso, where the identity of one inflection determines every other), but it is far from entirely unpredictable as well. And unlike Wurzel’s brief account of the fragment of Latin declension treated above, Bittner’s study also highlights the diachronic relevance of the directional implications he identifies. Given the ongoing historical tendency towards weak inflection as mentioned above and documented in Bittner (1996: 66f. and *passim*), the existence of an overarching PSC (2.5) suggests that any changes to the behaviour of individual lexemes in the direction of this weak default should affect certain parts of the paradigm before others, e.g. one would not expect the abandonment of an inherited strong past participle until after the abandonment of strong forms elsewhere (whose presence would imply its survival according to the PSC); and expectations along these lines are generally borne out (see e.g. Bittner 1996: 111, Dammel 2011: 197 on the directionality of ongoing developments of this kind).

At the same time, however, it is notable that where developments do take place in the contrary direction, whereby strong characteristics are *acquired* rather than lost in part of the paradigm of particular verbs, they do not always appear to proceed in accordance with the implicational schema above. For example, Dammel

observes (2011: 198) that while some verbs, such as *winken* ‘wave’, have taken on innovative strong inflection only in the past participle (giving optional *gewunken* alongside *gewinkt*), a different pattern of change is exemplified by *fragen* ‘ask’: this shows umlaut in the 2sg, 3sg present (*frägst, fragt*) in a range of regional and colloquial German varieties, while strong preterites and past subjunctives (*frug, früge*) are much rarer and a strong past participle is non-existent for this verb (others behaving similarly are *kaufen* ‘buy’, *fassen* ‘grasp’, cf. Nübling 2001: 465f.). The widespread adoption of such a pattern is unexpected if strong forms in the present imply strong forms elsewhere. Interestingly, a similar effect can be seen in Dutch, where an analogous strong/weak dichotomy obtains. Developments from strong towards weak have reliably affected the preterite before the past participle as in German; but developments in the other direction have *also* targeted the preterite first, giving e.g. *vragen* ‘ask’ the strong preterite *vroeg* alongside weak past participle *gevraagd* (verbs behaving similarly include *jagen* ‘hunt’ and *waaien* ‘blow’: Dammel, Nowak and Schmuck 2010: 345f., fn. 12), meaning that we observe a pair of opposed diachronic developments which cannot reflect the influence of a single implicational PSC. Such phenomena point to the existence of other contributory factors to the articulation of the inflectional system even when the case can generally be made for the validity of overarching implicational PSCs such as (2.5); in the Dutch case, Dammel et al. suggest (*ibid.*) that the past participle is relatively highly ‘cognitively anchored’ for reasons of frequency and accordingly is more likely to preserve its conservative inflectional behaviour, whether strong or weak.

2.2.3 Discussion

However this may be, the treatment of German verb inflection sketched out here, together with Wurzel’s account of the fragment of Latin nominal morphology above,

gives a representative picture of the Natural Morphology approach to the question of inflectional diversity. A comparison with the approach taken in the framework initiated by Carstairs-McCarthy, described above, allows us to identify differences between them which raise many of the central issues in play in a consideration of potential restrictions on this diversity and the related notion of inflectional class, upon which heterocclisis is based.

One salient difference, touched on already in the discussion of Carstairs-McCarthy's approach above, concerns the material whose distribution 'counts' for the purposes of the two general frameworks. For treatments of inflectional behaviour in the Paradigm Economy tradition, what is relevant is the distribution of inflectional *affixes* in particular: the notions of 'economy' and 'blur' relate specifically to the roles which can be attributed to affixes as the markers of distinct inflectional classes or otherwise as elsewhere items assigned by default. Inflectional classes in these terms are thus explicitly constituted only by differences in affix choice across the lexicon (or rather the part of speech in question). However, as noted by Blevins (2004: 53-55) among others, the methodological choice to restrict the scope of enquiry in this way allows for the possibility of ruling out of consideration a great deal of material which is in fact demonstrably relevant to paradigmatic organization. A clear example can be seen from the German case just discussed: even though affixal inflection clearly predominates in the morphology of the modern language, Bittner shows that vocalic alternations in the stem participate in the same complex of implications as affixes do, which suggests that to separate the two types of inflectional marking as a matter of principle (as is nonetheless seen in Carstairs-McCarthy 1991: 242-47) would be somewhat artificial.

Carstairs-McCarthy and others who make use of the same framework do acknowledge that it does not include all of inflection within its scope, rather arguing that it at least allows for the identification of principled restrictions with regard to a crucial inflectional domain (e.g. Carstairs 1983: 161; Enger 2007: 279f.). But more generally, the difference in focus shown by the two approaches above is connected with their somewhat different conceptions of how the theoretically possible inflectional ‘free-for-all’ alluded to above, in which every word’s paradigm could show its own individual combination of inflectional behaviour up to the limits of the morphological resources available, is ‘avoided’ by genuinely attested inflectional systems. The affixal economy framework in effect posits that the range of allomorphy existing in a language directly imposes a limit on the number of different inflectional patterns available; while the ultimate mechanism taken to underpin this limiting effect differs in successive versions of the framework, it operates essentially as an extrinsic constraint on the inflectional diversity of an entire part of speech, predisposing speakers to make changes if it is violated, and thus meaning that the number of combinations of inflectional behaviour shown by different lexemes can never increase above a certain modest limit to anything like the full number mathematically possible. Meanwhile, the Natural Morphology approach takes a different perspective on the question of the limits of inflectional diversity at issue here. As we have seen, this model recognizes inflectional allomorphy as intrinsically *unnatural*, from the point of view of universal morphological naturalness laid out by Mayerthaler, and ultimately as the result of compromise between the rival interests of various scales of naturalness both inside and outside the morphological domain: as such, it poses a challenge for speakers, which is mitigated by the development of implicational relationships, defaults and favoured reference forms and types of

exponence that help to reduce arbitrariness in the system and constitute a *local* standard of naturalness within it. But in such a framework – which emphasizes the view of an inflectional system as a complex of mutually reinforcing tendencies and preferences, rather than simply a combinatorial mechanism for distributing ‘given’ inflectional material over the lexicon – there is no reason to think that the presence of order rather than chaos in a language’s inflection need result from the existence of an overarching pressure to restrict the available patterns of behaviour to their logical minimum with regard to some single, privileged criterion, applied ‘top-down’ over the system as a whole.

Importantly for the present discussion, this difference in attitude with regard to the basis for inflectional organization goes hand in hand with a difference in attitude towards the nature of inflectional classes. The approach taken by Carstairs-McCarthy is founded not only on the observation that lexemes can be grouped according to their inflectional behaviour, but on the proposal that the number of these groups is itself a fact of structural importance to which the morphology is sensitive: inflectional classes in this sense operate as primitives in the system, which license the behaviour shown by the lexemes assigned to them. This is why the No Blur Principle is able to treat membership in a class as ‘part of the meaning of an inflectional affix’ (Carstairs-McCarthy 1994: 741): the affix reflects or identifies the class, which exists in its own right.

What is more, as illustrated above, for the economy principles at the centre of this tradition to be able to operate successfully they are taken to apply after a (potentially substantial) procedure of abstraction, in which numerous distinctions correlated with gender, phonology and more may be factored out to leave consolidated ‘macroclasses’. While this notion simply builds on a common analytical

practice which is often left implicit in the description of particular inflectional systems, it has been observed that it is hard to identify a principled way to rule out macroclasses that go beyond what is generally treated as intuitive and ‘innocuous’ in this way: thus Blevins (2004: 52) notes that while it seems reasonable to abstract away e.g. from the ‘hard/soft’ phonological dichotomy in Russian nominal inflection, there is no clear reason to *stop* consolidating the Russian declensions at the point where this is generally done. But it should also be recognized that, at whatever point such macroclasses are taken to be established, economy principles are left with little to say about formal relationships either between these macroclasses, or between the lower-level ‘unconsolidated’ classes which make up each one. That is, there is a single, privileged level of analysis, the macroclass level, at which the concept of inflectional class can be said to fulfil the role assigned to it in this framework: only here can inflectional classes be identified as fundamental entities in a sense relevant to the workings of the morphology.

Wurzel’s approach, meanwhile, demonstrates that it is possible to accord a very different status to inflectional classes. As we have seen, in *Natural Morphology* a given lexeme shows the pattern of inflectional behaviour that it does thanks to the particular way in which its lexically stored characteristics, which may or may not include specified inflectional forms beyond the basic form, interact with the set of defaults and implicational relationships which apply in common across the whole inflectional system. For example, the fact that a Russian noun whose (nom sg) basic form ends in *-a*, in the absence of any additional lexical specification, will inflect like *karta* ‘map’ above, is explained by this feature of its phonology in combination with the existence of a PSC which applies to basic forms with this shape; meanwhile, the inflection of *zakon* is explained directly by its ending in a ‘hard’ consonant, in an

analogous way. From this point of view, inflectional class labels (such as ‘second declension’ for *karta*) merely serve as shorthand for the identification of different patterns of behaviour, and play no explanatory role: that is, the claim that *karta* belongs to the Russian ‘second declension’ is not envisaged to provide an account of its inflectional behaviour at some deeper explanatory level, but merely restates the fact that *karta* belongs to the group of nouns that inflect in that way by virtue of being subject to the relevant PSC. The notion of inflectional class is not a primitive in this approach but a way of generalizing over lexemes with regard to their inflectional behaviour, which is itself accounted for by other means.

Wurzel’s conception (1984: 66f.) of what it means for two lexemes to belong in the *same* inflectional class appears to emerge straightforwardly from this ‘bottom-up’ as opposed to ‘top-down’²¹ approach to inflection. For him, two lexemes share a class only if their inflectional behaviour agrees in *all* respects: a single difference in inflection anywhere in the paradigm is sufficient to assign lexemes to separate classes – with the sole exception that formal distinctions in the *basic* form (whatever this is taken to be) are not considered relevant to inflectional class assignment by definition.²² This maximally fine-grained treatment allows even very small groups of lexemes to be identified as classes in their own right, and indeed Wurzel posits no lower limit to the possible size of an inflectional class; two lexemes may constitute a class, or indeed even a single lexeme, at least with the proviso that its behaviour actually arises from the implicational structure that other lexemes participate in and is not simply lexically listed as a special case (Wurzel 1990: 212). This strict understanding of what constitutes distinctive classes evidently has the potential to identify them in great numbers, especially as the relevant differences are not limited

²¹ For these terms for this opposition cf. Dammel (2011: 23), ‘von unten’/‘von oben’.

²² ‘Einheitliche Flexionsklasse heißt nichts anderes als einheitliche Bildung der abgeleiteten Flexionsformen’ (Wurzel 1984: 68).

to affix choice but can involve stem behaviour as well. This is not in itself problematic: as the different patterns of behaviour shown are themselves the point directly at issue, and nothing rides on the precise number of classes distinguished at the analytical level of interest, there is no rationale for aiming to distill the attested variation into a reduced number of higher-level classes, as we have seen is essential for the success of affixal economy approaches.

Prominent recent work directly addressing the question of inflectional diversity, which takes as its material a wide range of typological data illustrating the different types of organizational structure available in inflection across the world's languages, endorses the implicit Wurzelian position that there is no good reason to identify and privilege a special 'macroclass' level (in Carstairs-McCarthy's sense) in the treatment of inflectional systems – instead casting doubt on the fundamental status of proposed principles of affix economy and on their power to explain the layout of specific systems or the existence of predictability in inflectional behaviour in general. The empirical validity of the No Blur Principle, and its successor Vocabular Clarity, has been challenged from various corners on the grounds of its failure to account for data from particular languages (e.g. Sanskrit: Stump 2005b, 2006; Polish: Halle and Marantz 2008; Nuer: Baerman 2012; Võro (South Estonian): Baerman 2014), and despite defences of the approach (e.g. Enger 2007, especially pp. 284-86; Carstairs-McCarthy 2014), the existence of serious counterexamples seems inescapable. Thus Stump and Finkel (2013: 103-7), reporting their earlier findings (Finkel and Stump 2009, following Jakobi 1990) on the organization of verb morphology in Fur, point out that they imply severe violations of both Paradigm Economy and No Blur, even while the analytical strategy of factoring out non-affixal

distinctions between inflectional patterns (which here involve tone) already substantially underrepresents the genuine diversity involved in the system.

But perhaps just as telling as the existence of individual counterexamples is the observation, made recently by Ackerman and Malouf 2015, that the empirical effects attributed to No Blur need not in fact be taken to arise from the existence of any such morphological principle, or from the operation of the more fundamental Principle of Contrast supposed to underlie it – a point that they illustrate by means of a simple model of inflectional acquisition, which shows how scenarios compatible with No Blur can come about thanks simply to the cumulative effects of analogical reasoning, even of the most elementary kind. While the details of the iterated learning model (Kirby and Hurford 2002) employed in Ackerman and Malouf’s demonstration will not be dwelt on here, the crucial point is that a simulation of the learning process applied to a constructed language over successive ‘generations’ of speakers, which builds in very few stipulations about how that learning progresses (and in particular gives no role to the Principle of Contrast), can nonetheless lead from very highly ‘blurred’ starting conditions to the kind of orderliness and (relative) lack of blur observed in attested inflectional systems.

As in earlier publications (Ackerman, Blevins and Malouf 2009; Ackerman and Malouf 2013), they begin by raising what can be considered the crucial question facing speakers of a language with a complex inflectional system (and thus, from one point of view, the crucial question about the organization of inflection cross-linguistically), namely the *Paradigm Cell Filling Problem*, stated in its most recent form as follows (Ackerman and Malouf 2015: 5): ‘Given exposure to an inflected wordform of a novel lexeme, what licenses reliable inferences about the other wordforms in its inflectional family?’ This draws attention to the fact that speakers

are regularly called upon to produce (or interpret) inflected forms which they have never encountered before and must therefore deal with on the basis of their prior, necessarily partial, knowledge of the inflectional system. Stripped back to its simplest terms, then, this system as it is known to any given speaker consists of a set of lexical items each associated with a set of paradigm cells, some of which are already occupied by previously encountered forms while others are empty and need to be predicted.

For a speaker wishing to predict the contents of an empty cell C in a given lexeme's paradigm, one (minimally sophisticated) type of analogical reasoning might be to identify a single 'predictor' cell in that lexeme's paradigm for which the correct form *is* known, survey all the different known inflectional patterns which share its behaviour in that predictor cell, and then make use of the inflection 'prescribed' for cell C by the greatest number of those patterns. This is the strategy assigned to a simulated speaker by Ackerman and Malouf's model (2015: 7): in each generation, precisely one inflectional form is set at random to be unknown to the speaker and thus to require prediction in this way, on the basis of full knowledge of the entire remainder of the system. Crucially, the form predicted then becomes 'correct' for the purposes of the next generation, to be supplied as prior knowledge and thus available to direct the course of the language's future development; the earlier form, supposing that it was different from that predicted, is lost.

Evidently, the precise trajectory of a system's evolution over the generations when it is subjected to this model will depend both on its original characteristics and on random chance. But what Ackerman and Malouf are able to show is that, even when the original inflectional systems are designed to be as blurred as possible,²³

²³ Specifically, a system was set up in which a lexeme's paradigm had eight cells which could each be filled by one of three 'allomorphs', giving a total of 3^8 (= 6561) potential combinations of inflectional

repeated simulations result sooner or later in stable inflectional systems (defined as existing for 25 generations without any change, i.e. without any ‘incorrect’ predictions being made) of which the majority obey the No Blur Principle. In many cases, this is for the trivial reason that inflectional allomorphy has been lost altogether in the course of the simulation; however, from their sample of 500 languages that do retain two or more inflectional classes after stabilizing in the relevant sense, they report (p. 8) that 56.7% still satisfy No Blur. Given the many features of this model which evidently distance it from any genuine linguistic situation, it would be a mistake to place too much weight on this precise figure. However, the essential point is established that the absence of blur can arise as an *emergent* property from the operation of a simple analogical procedure which does not in any way stipulate it as a desired result in advance: in fact, as Ackerman and Malouf point out, the fact that many but by no means all such simulated languages accord with the No Blur Principle is reminiscent of precisely what we seem to observe across the genuine inflectional systems which have been investigated with this question in mind.

From a broader perspective, Ackerman and Malouf’s aim in this paper is to suggest that No Blur effects, and by extension other affixal economy effects, are in fact merely side effects or facets of a more general phenomenon which they also investigate in their previous work (Ackerman, Blevins and Malouf 2009; Ackerman and Malouf 2013), namely the pressure for inflectional systems to show low *conditional entropy* (i.e. relatively high predictability of individual forms) in order to be learnable and usable by speakers. While a precise mathematical characterization

behaviour over the paradigm. Each simulation took as its initial state an inflectional system consisting of 100 lexemes whose inflectional behaviour was selected at random from those 6561 combinations, meaning in practice that within a given simulation every lexeme’s inflection was most likely unique. In such a scenario, no allomorph can be considered a ‘class-identifier’.

of conditional entropy in this sense would take us too far afield here,²⁴ its value, as calculated for a given inflectional system, is designed to capture the general level of uncertainty encountered by a speaker when faced with the problem of supplying an unknown inflected form in that system, such that the lower the value, the less severe this uncertainty: an average conditional entropy of zero bits means that any inflected form of any lexeme is entirely predictable on the basis of any other of its inflected forms, as in the situation presented for the fragment of Burmeso inflection provided above (section 2.2.1).²⁵

It might be expected that average conditional entropy would be found to vary widely from language to language, in view of the substantial variation observed cross-linguistically with regard to such measures of inflectional complexity as paradigm size and inflectional diversity across the lexicon (subsumed as E[enumerative]-complexity by Ackerman and Malouf 2013: 429). However, in investigating this expectation using a ten-member sample of typologically diverse inflected languages, the authors find that despite the ‘dramatically differing degrees of morphological E-complexity’ (p. 444) of the inflectional systems selected, their calculated conditional entropy is uniformly low: in nine out of the ten languages, its value is at or below 0.75, which is equivalent to saying that prior knowledge of the form in one cell leaves a chance greater than one in two of selecting the correct form in another – even if the system as a whole contains dozens or hundreds of inflectional classes. That is, at least within this sample, those inflectional systems which appear daunting from the point of view of E-complexity do not in general pose

²⁴ For a mathematical description of this and related concepts originating in information theory (Shannon 1948), cf. Ackerman, Blevins and Malouf (2009: 62-68); Ackerman and Malouf (2013: 437-42).

²⁵ Note that Ackerman and Malouf’s (2013: 443) Table 3 mistakenly gives the average conditional entropy for Burmeso as 1.000 bits (in fact, the labels for the final two columns appear to have been switched, meaning that values are misidentified for all languages): the correct value of zero bits is stated in the text on p. 445.

greater challenges to speakers seeking to predict unknown forms than those whose E-complexity is lower.

In the light of this striking result, they conjecture that the crucial restriction on morphological systems concerns not the E-complexity, but rather the average value for conditional entropy, thus ruling out the possibility of systems in which it is overly difficult to extrapolate from known to unknown forms, but without also imposing constraints on how this level of inter-cell predictability is to be achieved. Crucially, they note that conformance with affixal economy (however defined) is just one characteristic a system might possess which would help to ensure that it showed low conditional entropy, and this being the case, it should be no surprise to find that not all languages appear to observe affixal economy as would be predicted if it were (or were directly produced by) a fundamental principle: the same desirable low-entropy state of affairs could be achieved either by the restriction of inflectional diversity and morphosyntactic distinctions, or by the existence of especially reliable principal parts and implicational patterns within the paradigm, which would have the power to take practically all the ‘guesswork’ out of the analogical reasoning used in the prediction of unknown forms. While they present these possibilities as alternative *strategies* for bringing about low entropy, it makes equal sense to treat them all as alternative *by-products* of low entropy, which can be taken as the feature directly favoured by learnability considerations (pp. 453f.) and is the common feature observed across the languages sampled – as across the model systems resulting from the evolutionary learning simulations discussed above.

Related work in Stump and Finkel 2013, which likewise treats the nature of inflectional predictability and the cell-filling problem but is more narrowly geared towards the study of the notion of principal parts, both endorses this finding on

conditional entropy on the basis of independent data (cf. especially chapters 5 and 10) and reports other instances of striking cross-linguistic similarity bringing together typologically diverse inflectional systems. For example, they find that the average ‘cell predictor number’, i.e. the number of principal parts required on average to determine the correct form for an unknown cell, ranges only from 1.00 to 1.09 over nine out of ten systems of verb inflection selected to represent a diverse range of inflected languages (while the value for the tenth, Kwerba, is much lower only because some cells of the paradigm show no inflectional allomorphy whatsoever: Stump and Finkel 2013: 60f., 78).

Such typological investigations of the nature and effects of inflectional implication over the paradigm and over the lexicon in effect follow along the lines laid down by Wurzel and the Natural Morphology approach, in which particular attention is given to the relationships between cells in the paradigm. As Ackerman and Malouf point out, they in fact expand on the notion of reference forms (principal parts) by observing that in complex morphological systems practically all forms in the paradigm have some diagnostic value: ‘we can recognize predictiveness and predictability as gradient concepts that follow more generally from the organization of the paradigm’ (Ackerman and Malouf 2013: 453).

Accordingly, as observed above, these treatments agree with Wurzel in not granting any crucial structural status to an abstract ‘macroclass’ level, and instead preferring to focus on more concrete data. But this is not to say that they, or others interested in accounting for inflectional diversity, entirely endorse Wurzel’s own notably strict definition of inflectional classes. One point, which is of course of particular relevance to the present work, concerns the behaviour of nouns which are here taken to display heteroclisis, such as Czech *pramen* ‘spring’. As was seen in

section 1.1, this noun patterns with the so-called ‘soft masculine’ declension of *pokoj* ‘room’ in the singular and with the ‘hard masculine’ declension of *most* ‘bridge’ in the plural. To treat this particular case in Wurzel’s terms: since the morphosyntactic division between singular and plural is one which also plays a structural role elsewhere in Czech nominal paradigms²⁶ and thus *pramen* cannot be disregarded as simply irregular (cf. Wurzel 1990: 212 as discussed above), the fact that its inflectional behaviour across the paradigm differs from that of both *pokoj* and *most* necessarily establishes that it belongs to a different inflectional class from them both (cf. Wurzel’s 1984: 69 treatment of the inflection of German *Staat*). However, to identify *pramen* in this way as the representative of an inflectional class in its own right, existing alongside those of *pokoj* and *most*, seems to gloss over an important morphological fact, as Stump and Finkel observe (2013: 181): namely, it is part of a Czech speaker’s morphological knowledge that the existence of a strict distinction between the declensions of *pokoj* and *most* is well-established on the basis of a large number of lexemes, while the behaviour of *pramen* is highly exceptional in challenging this distinction and thus compromising otherwise strong generalizations about the inflection of masculine nouns. As a result, Stump and Finkel question whether it is appropriate to grant ‘autonomy’ (p. 175) to the behaviour of *pramen* (as would Wurzel), rather than instead preferring an ‘annexation’ approach which assimilates it to the classes it seems to straddle, i.e. recognizes that its behaviour is in some intuitive sense ‘parasitic’ on the existence of these two substantial classes which between them subsume the inflection of most masculine nouns in Czech. They conclude that an analysis which aims to capture what is morphologically significant to the language user will choose the second option in this case, while recognizing

²⁶ Cf. for example the inflection of *kuře* ‘chick’ (gen sg *kuřet-e*, nom pl *kuřat-a*) which shows weak stem suppletion correlated with the number distinction (Heim 1982).

that such analyses become less appropriate as the number of lexemes involved grows (Stump and Finkel 2013: 181, 366).

This ‘annexation’ approach to heteroclisis, which treats the behaviour of heteroclitite nouns as secondary to the prior existence of more ‘genuine’ classes, is not specific to Stump and Finkel 2013 in recent morphological research. It is the line taken, for example, by Brown and Hippisley 2012 in a study devoted to the organization of morphological and especially inflectional behaviour over the lexicon and the paradigm, which happens to take as its primary material the Russian system of nominal inflection as introduced above. They note the existence of certain nouns (such as *oblako* ‘cloud’, p. 149) whose inflectional behaviour follows that of fourth declension *slovo* ‘word’ throughout the paradigm with the exception of the genitive plural, which takes the ending *-ov* proper to the first declension as exemplified by *zakon* ‘law’ (giving *oblakov* instead of **oblak* etc.); meanwhile, *soldat* ‘soldier’ (pp. 35f.) behaves like first declension *zakon* with the exception that its genitive plural takes a zero ending (*soldat* rather than **soldatov*), as would be expected of nouns belonging to the second and fourth declensions. Rather than taking the groups represented by *oblako* and *soldat* to constitute separate inflectional classes of their own, Brown and Hippisley (*ibid.*) identify them as exceptionally combining the inflectional behaviour proper to two different inflectional classes, an analysis which in their model means that such words exceptionally inherit inflection in different paradigm cells from different inflection class ‘nodes’. Likewise, Corbett (2009a: 6), as part of a discussion of *canonical* inflectional classes to which we will return below, draws attention to the existence in Russian of a noun *put* ‘way’, which inflects throughout its paradigm exactly like third declension *noč* ‘night’ above with the sole exception of its instrumental singular *put’om*, which he observes to pattern

with the first (in fact this behaviour is proper to the fourth as well); accordingly, he presents the paradigm of *put'* as bringing together morphology from distinct inflectional classes, rather than embodying one in its own right. Indeed, Corbett refers to *put'* as a heteroclite noun belonging to the third declension, albeit with a borrowed form from elsewhere, just as Brown and Hippisley (2012: 35, 149) treat *soldat* and *oblako* as belonging to the first and fourth declensions respectively.

It is notable that this analytical move, of treating *put'*, *soldat* and *oblako* as (aberrant) members of otherwise well-motivated inflectional classes, is made more straightforward by the fact that each of them diverges from the regular behaviour of the respective classes only in a small part of the paradigm; it is less obviously attractive to identify *pramen* as belonging to a particular established inflectional class in this way even on the basis of an 'annexation' approach which denies it a class of its own. But the fact of taking this step even for *oblako* etc. implicitly distances the relevant conception of inflectional classes from Wurzel's own in an even more fundamental way, which can be seen to touch on a somewhat problematic aspect of his view.

Recall that for Wurzel, lexemes belonging to a single inflectional class cannot be at all distinct in inflectional terms anywhere in the paradigm (outside the basic form). As observed by Dammel (2011: 23), this is '[d]ie "härteste" Definition, die zu finden war', and evidently rules out the possibility of placing heteroclite *put'* in a class with regular *noč'*. But as she goes on to point out, the strictness of Wurzel's position has more general ramifications, in that it entails a binary notion of inflectional sameness: either two lexemes belong together, or they belong apart, and the question of *relative* similarity between the inflectional behaviour of different lexemes is accordingly glossed over. The same point can be raised at the level of

classes themselves as they are treated in Wurzel's approach: given that the inflectional behaviour of lexemes is implied directly by their extramorphological properties and/or reference forms, he is somewhat less interested in the levels and patterns of formal relatedness that may be present between the classes that lexemes thereby fall into (Dammel 2011: 23-28, 75f.).²⁷

Dammel dwells specifically on Wurzel's conception of inflectional class, and on its consequences for the line taken by his research programme, precisely because its restrictiveness is extreme within morphological scholarship. While it is not unusual to find authors apparently endorsing strict definitions of inflectional class similar to Wurzel's, in practice they do not as a rule limit their attention to the discussion of entirely homogeneous classes: for example, as seen in the introduction to the present work, Aronoff (1994: 64) states that 'an inflectional class is a set of lexemes whose members each select *the same set* of inflectional realizations' (emphasis mine), but this identification is not borne out in his subsequent discussion of Latin nominal inflection (pp. 79-88), in which he clearly has no intention of dividing up the nominal lexicon as finely as possible (as is attempted e.g. by Dressler 2002). Others are more explicit in eschewing a narrow definition: 'An inflection class is defined as a group of words that inflect *in the same or a similar fashion*' (Enger 1998: 140; emphasis mine).

The practice of investigating such classes which may nonetheless contain internal diversity does not betray a lack of respect for the data, but instead acknowledges the concept of inflectional similarity as a gradient phenomenon: it can make sense to group lexemes in a class in recognition that they behave *relatively*

²⁷ Such formal relatedness may not be evident from lexically listed properties themselves: for example, the fact that (according to Wurzel 1984: 120 and Table 2.9 above) Latin *ignis* 'fire' and *auris* 'ear' inflect alike in all but one cell is not reflected in their reference forms, which are respectively abl sg *ignī* and abl sg *aure*, acc pl *aurīs*.

similarly in comparison with others, rather than identically. And in fact, despite Wurzel's strictures, a reliance on relative similarity rather than sameness in the identification of classes is hard to avoid. To take a simple example: in the Russian second declension, exemplified by *karta* 'map' above, the genitive plural is usually marked by a 'zero ending', i.e. is distinguished from the nominative singular by the absence of final *-a* (here giving *kart*). Certain consonantal clusters which would thus surface in final position in the genitive plural are broken up by a 'buffer vowel' for phonological reasons: e.g. *sosna* 'pine' > gen pl *sosen*, *derevn'a* 'village' > gen pl *dereven'*. However, it is a further characteristic of many nouns whose nominative singular ends in *-n'a* that this consonant is depalatalized in the genitive plural form: thus *pesn'a* 'song' > *pesen*, not **pesen'* (Wade 1992: 70f.). In Wurzel's terms, this disagreement in stem behaviour between *derevn'a* and *pesn'a* allows us to identify two separate classes here. However, in practice the fact that *derevn'a* and *pesn'a* behave identically with respect to their inflectional endings and in all other respects means that this minor distinction is generally overlooked for the purposes of classification, as it is in the presentation of Russian noun morphology above. Disregard for such points may seem to be equivalent to a covert reintroduction of Carstairs-McCarthy's concentration on affix behaviour at the expense of stem behaviour. However, the conflation of *derevn'a* and *pesn'a* in a single class for the purposes of an analysis of Russian inflection makes no claim about the nature of inflection in general – instead it simply acknowledges the fact that slight stem alternations of this kind are rare and of relatively little importance for the organization of Russian nominal inflection, interacting little with any other features of the system, while differences in endings are ubiquitous and play a crucial structural role, fundamentally articulating the inflectional system in a way that stem

behaviour does not. Allowing for similarity rather than sameness in a conception of inflectional classes thus enables us to focus attention more easily on what can be considered the most significant distinguishing factors in inflection over the lexicon in a given language, albeit at the expense of opening the door to differences of interpretation on what constitutes this significance and where ‘appropriate’ levels of similarity should be taken to lie (though note that inflectional sameness itself is hardly an uncontroversial notion: Enger 1998: 28).

In effect, then, a loosening of the notion of inflectional classes is tantamount to the recognition that the most significant organizational principles of the system, from the point of view of the analyst but also of the language user, are likely to be found at a level of generality higher than the very lowest: this is a position that can be upheld without requiring a belief in the fundamental status of a particular, abstract macroclass level in Carstairs-McCarthy’s terms. Indeed, as observed by Dammel (2011: 25, fn. 17), even Wurzel endorses the value of a looser notion of classes when he effectively treats the German ‘strong verbs’ as comprising a single class (1984: 72, cf. 172), despite the diversity treated above as well as questions of stem ablaut that have not been touched on here. This is the basic intuition behind the construction of inheritance hierarchies (Haspelmath 2002: 125-30) as a tool for analysis and/or a proposed reflection of speaker knowledge: these are designed to capture the fact that different *levels of similarity* can be identified, implying classes at various levels of generality, which are all worthy of attention and may all be presumed to be psychologically real to the speaker. In fact, Dressler adapts the Natural Morphology approach to inflectional classes in a way which allows for the construction of such hierarchies (e.g. Dressler 2002, 2003), thereby enabling comparison of the relative ‘depth’ of such hierarchies for typological purposes (Dressler et al. 2006): notably,

the fact that homogeneous inflectional behaviour in his model assigns lexemes to inflectional ‘microclasses’ or even ‘minimicroclasses’, with ‘classes’ potentially many steps higher in the structure and ‘macroclasses’ (to be distinguished from Carstairs-McCarthy’s concept) yet higher, is indicative of an acknowledgement that often the view of inflectional phenomena at the finest-grained level is less illuminating and linguistically significant than one at a higher level which combines specificity with generality.

In the context of the present work in particular, this gradient and hierarchical approach to the question of similarity within inflectional classes is valuable because it allows for the observation of instances of heteroclisis involving patterns of inflectional behaviour at levels of generality above the very lowest. Indeed, the relevance of the notion of heteroclisis to classes which would not meet Wurzel’s criterion of total internal homogeneity has been assumed above – for example, the claim made in section 2.1 (cf. Table 2.3) that Romanian *a fi* ‘be’ shows heteroclisis between the first conjugation in the imperfect and the third conjugation in the preterite does not rely on the claim that either of these conjugations constitutes a microclass in Dressler’s sense. This is the line also taken by Dressler et al. (2006: 53, 60) with reference to German verbs: while they assign weak verbs to Macroclass I and the majority of verbs with strong characteristics to Macroclass II, each divided into lower-order classes, they posit no class at any level for verbs such as *melken* (-*gemolken*) with strong past participles only, instead implicitly ‘annexing’ the behaviour of such verbs to that of the two macroclasses which it straddles. Meanwhile, the illustration given above (cf. Table 2.2) of the interaction between heteroclisis and tense-deponency in the group of Gothic verbs known as the ‘preterite-presents’ also dealt with heteroclisis at a level applying to strong and weak

verbs in general. Not to do so in this case, because of the internal diversity of both the strong and the weak verbs in Gothic and the corresponding differences of detail in the behaviour of the fourteen preterite-presents (cf. Tanaka 2011: 12f.), would be to miss a striking generalization: ultimately it is useful to be able to say that a single phenomenon is in question here.

This acknowledgement that it can be beneficial to address the behaviour of inflectional classes at different levels of generality, rather than identifying and then privileging a particular level cross-linguistically on grounds of principle, points to the utility more generally of taking each inflectional system ‘on its own merits’. The complexity of the question of inflectional diversity and its potential limits, of which the extensive discussion above has nonetheless only been able to give a first impression, may seem to suggest that ultimately there is no correct standpoint to take up in order to gain a full understanding of the organization of inflection, even though we never find ourselves in the theoretically possible situation of total inflectional chaos invoked by Carstairs 1983; the work of Ackerman and Malouf, for example, seems to cast doubt on the idea that a single principle of structural organization does, or should be expected to, account for the relatively ‘low entropy’ observed in attested inflectional systems, whether those systems are easy or difficult to capture from the point of view of the analyst. One upshot of this is that it is desirable to build a descriptive framework for inflectional morphology which does not impose its own inbuilt structure on the treatment of the particular system it investigates, but starts from the assumption that any inflectional behaviour may *in principle* be observed and that, on that basis, cross-linguistic research is concerned with uncovering what is and is not observed in reality: this is the aim of the framework of Canonical Typology, which attempts to lay out the abstract theoretical space in which

variability can operate both within and across languages, and I return to it immediately below in section 2.3. However, the fact of low entropy, and thus the existence of principled presences and absences among the inflectional combinations potentially available in every system, means that with regard to any one of those systems there will be regularities to identify, whatever their nature might be; one of the tasks of the morphologist is to determine where best to ‘start from’ in order to address these regularities.

As stated above (section 2.1), and as exemplified in the brief illustration of heterocclisis found alongside stem suppletion in the paradigm of Romanian *a fi*, my own treatment of heterocclisis in Latin and Romance as presented in the following chapters takes as its ‘starting point’ the level of generality in the inflectional organization of the lexicon, and the source of historical continuity in exposition, which are provided by the traditional four conjugations that have long been recognized for the Latin verb system (in fact since at least the third century AD).²⁸ These four conjugations are by no means homogeneous classes; nor, indeed, are they macroclasses in Carstairs-McCarthy’s sense, in that the internal variation they encompass cannot be motivated by or reduced to the effects of independent factors. Rather, I take them to stand at an appropriate level to capture the most salient and linguistically significant generalizations about inflectional diversity in the Latin verb. As will be touched on again in section 3.1, the conjugations partition the lexicon almost without remainder, are internally coherent with regard to important features such as the selection of much tense and mood marking, and are characterized by the presence of distinctive theme vowels (relevant for derivation as well as inflection) whose distribution over the paradigm, and whose interaction with the inflectional

²⁸ Cf. Taylor (1991: 88f.)

material they accompany, are conjugation-specific. Moreover, while the conjugational distinction is primarily relevant to a particular section (in fact constituting the majority) of the verb paradigm, featuring the so-called *infectum* stem, there are also certain relationships to be observed, characterizing the different conjugations with different levels of reliability, between this and the verb's other stems (cf. Aronoff 1994: 49 and sections 3.3, 3.4 below). It is for reasons including these that the four conjugations as traditionally identified are taken as the focus of attention, as concerns the organization of inflectional diversity over the lexicon, not only in the present work but in practically all analytical works which have treated Latin verb morphology, e.g. Matthews 1972, Carstairs 1987, Aronoff 1994. For my purposes it is also relevant that (the continuants of) these conjugations remain central to the characterization of a given verb's inflectional behaviour in the Romance varieties I go on to examine, and indeed distinctly Romance innovations can be identified in the development of the verb system which are limited to particular conjugations, e.g. the 'inflectional augments' seen in the first and fourth conjugations in Romanian and Romansh, cf. Meul 2013 and chapters 4 and 5 (especially section 5.4) below. This longstanding structural importance of the traditional conjugations in articulating inflectional behaviour over the lexicon in Latin and Romance is one reason why it is rewarding to examine instances of heteroclisis which compromise the inflectional norms identifiable at this rather high level.

2.3 Heterocclisis in Canonical Typology

The so-called ‘canonical’ approach to the typology of inflectional morphology which I present here, as seen e.g. in Baerman, Brown and Corbett (2005: 27-35), Spencer 2007, Thornton 2011, Palancar 2012, and Corbett 2007a, 2009a, 2015, forms part of a wider typological programme, which has also been applied in the domains of phonology (e.g. Hyman 2009) and syntax (e.g. Comrie 2003, Nikolaeva and Spencer 2008). As just stated, the intent behind this programme is to chart an abstract cross-linguistic framework in which to view linguistic data: this is carried out by recognizing dimensions along which there is the potential for linguistic variability (within and across languages), and then identifying the logical extreme or extremes of that variability along each dimension in order to produce a theoretical space within which genuinely attested phenomena can be situated. These extremes may, but need never be encountered in practice; rather, they represent the endpoints *in principle* of particular scales, in relation to which genuine instances can be considered. Corbett (2015: 153) thus offers the analogy of the temperature scale, which is fixed with absolute zero as its cardinal point and allows for comparison of all other temperature data against that point, without thereby implying that a temperature of absolute zero has been, or ever could be, observed in the physical world.

This approach to linguistic phenomena has the advantage of making explicit the fact that the descriptive notions that linguists are interested in using, whether in order to make cross-linguistic comparisons or to weigh up data from a single language, often implicate several such dimensions of variability at once, and that disagreements about whether a particular datum counts as an instance of the phenomenon of interest are often based on the question of how crucial one or other

of these dimensions may be, or where to ‘draw the line’ along it: for example, whether something is an instance of *agreement* (for gender) if the gender in question is not overtly present on the controller (for example, in Russian *ja pisal* (m) / *ja pisala* (f) ‘I wrote’ (Brown and Chumakina 2012, after Corbett 2003)). Accordingly, it should be possible to characterize what would constitute *canonical* examples of particular linguistic phenomena that we wish to label by identifying the relevant dimensions of variability and establishing where they ‘converge’ on the point of interest. With regard to inflectional morphology, it is possible to define criteria for *canonical inflection*, which will be presented below: crucially, it is thus also possible to identify different attested inflectional phenomena as non-canonical in different ways. The concept of heteroclisis has been touched on in several presentations of the Canonical Typology approach which, although their focus lies elsewhere, have recognized that heteroclisis should be recognized with other inflectional phenomena such as suppletion and deponency as non-canonical, and in particular as non-canonical features entailing ‘lexical splits’ in the paradigm (Corbett 2015); in turn, however, we can talk about the features making up canonical suppletion, deponency, or indeed heteroclisis.

In a system making use of inflectional morphology, lexemes take different forms which realize different grammatical functions: the existence of these overt differences is what leads us to identify the array of inflectional distinctions in the paradigm that we do in fact identify. In an inflectional system at the logical limit, then – the inflectional system that would best deserve the name – the relationship in both directions between the inflected forms of lexemes and the lexical and grammatical functions which they possess would be unique in individual cases and consistent over the lexicon. This is captured in Table 2.11, by now familiar in the

literature on canonical typology: the digits placed in brackets are referred to below in the text.

Table 2.11 The properties of canonical inflection (after Corbett 2009a: 2)

	Comparison across <i>cells</i> of a lexeme	Comparison across <i>lexemes</i>
Composition/structure (morphotactics)	same (1)	same (4)
Lexical material (\approx stem)	same (2)	different (5)
Inflectional material (\approx inflection)	different (3)	same (6)
Outcome	different	different

The central column in this table states simply that in a canonical inflectional system, the lexical material within a single lexeme will remain the same (in morphological terms), while the inflectional material will differ from cell to cell as it encodes different inflectional features; meanwhile, the way in which lexical and inflectional material interacts (suffixation, fused exponence, periphrasis etc.) will remain constant. Meanwhile, comparison across lexemes in such a system, as detailed in the right-hand column, will reveal that the lexical material is different in every case, while the same inflectional features will be represented by the same inflectional material regardless of the lexeme involved, and the morphological structure of inflected forms will remain constant throughout. The system which results is one in which every inflected form of every lexeme is distinct using the minimum of inflectional machinery.

This canonical view of how inflection could in principle operate, in the ideal case, allows us to identify and compare the ways in which genuine inflectional behaviour deviates from it: a wide array of inflectional phenomena can be recognized as embodying a deviation from one or another of the characteristics of canonical inflection. Thus consider a deviation from canonical behaviour in (2): this would

imply stem allomorphy or, in the extreme, suppletion (Corbett 2007a). Syncretism is deviation from the canonical behaviour required in (3) (Baerman, Brown and Corbett 2005). Meanwhile, as noted by Corbett 2007b, one type of deviation from (6) is deponency, in which a lexeme bears inflectional material in a given cell which ‘belongs’, on the basis of comparison across lexemes, to another cell, thus appearing to encode the wrong morphosyntactic properties.

Importantly, however, each of these types of non-canonical behaviour is in turn susceptible itself to treatment according to the canonical approach. Thus, given that the non-canonical phenomenon of suppletion exists, it is possible to investigate what would constitute the ultimate instance of the concept, thereby identifying various ways in which genuine morphological behaviour might match with or diverge from this. For suppletion, Corbett 2007a provides fourteen criteria for canonicity, i.e. fourteen dimensions along which attested instances may differ from the canonical case: for example, according to Criterion 4 (p. 17), *more variants* > *fewer variants* of the lexical stem (where ‘>’ represents ‘is more canonical than’), as the constancy of the lexical semantics across the paradigm is encoded more and more poorly the more different stem forms are required by the inflectional system. Likewise, according to Criterion 2 (p. 15), *full* > *partial* suppletion, because again the lexical material required by the paradigm more thoroughly ‘lets down’ the semantics the more drastically it fails to cohere over the paradigm. Meanwhile, Criterion 12 (p. 27), *unique* > *nonunique* with regard to patterns of alternation shown over the paradigm, captures the sense that the more lexemes can be found to show the same pattern, the more transparent and less irregular the particular instance of suppletion is.

Heteroclisis figures along with these better-known inflectional traits among the types of non-canonical behaviour discussed by researchers in the canonical typology framework. For example, Spencer 2007, in work based on that of Corbett, treats heteroclisis, syncretism and deponency as three varieties of ‘intra-class’ morphological mismatch. However, relatively little attention has been paid to the precise way in which heteroclisis fits into the basic picture arising from Table 2.11. Thus, for instance, as well as identifying deponency as a violation of the aspect of canonical inflectional behaviour which I have noted above as (6), Corbett (2007b: 27; 2015, 152) also lists both ‘inflectional classes’ and ‘heteroclisis’ under the same heading. But this is somewhat misleading, as what belongs in this slot, alongside deponency, is *inflectional allomorphy*, i.e. purely formal variation in inflectional realization, which does not involve the ‘misuse’ of an inflection proper to a different cell but still distances the inflectional system from canonicity in the same broad terms. The existence of inflectional classes is itself predicated on the existence of inflectional allomorphy; and in turn, heteroclisis is only possible given the existence of inflectional classes. Heteroclisis can thus be considered a type of non-canonical behaviour of a higher order than those seen so far.

The point made here is in fact brought out tacitly in Corbett 2009a, which treats canonical inflectional classes. Criteria identifying canonicity here include, for example, the central factor that in the canonical situation every form belonging to one class should differ in its inflectional material from the corresponding form belonging to a different class (Criterion 1, p. 4), because this makes the distinction between the two classes as stark as possible, and the difference between the inflectional behaviour of the two classes was what led us to identify inflectional classes here in the first place: formal differences everywhere that they are possible

take the definition to its logical limit. Corbett notes that the strict division ideally found between inflectional classes may be compromised by the existence of ‘a small number of items showing combinations of forms from other classes [which] can be treated as heteroclites’ (p. 6) and which thus themselves represent a deviation from canonicity at this higher level.

It is interesting that he does not take the further step of considering that heterocclisis itself may be susceptible to a canonical treatment, like inflectional classes and the other morphological phenomena laid out above. However, such a move evidently follows from the logic of the overall approach. For example, rather than simply stipulating that heteroclites will be few in number in comparison with the inflectional classes they are related to, one should treat *unique* > *nonunique* as a criterion for canonicity in heterocclisis as it is in suppletion: the fewer lexemes there are which use a given pattern, the less similar they are to an inflectional class of their own. Likewise, the correlation of heterocclite behaviour with possession of a particular lexical feature, such as a gender value (cf. the abstract schema discussed as Noyer’s ‘systematic mixed inflection’ in section 2.4 below and illustrated in Table 2.13, and the concrete case of the Gothic *i*-stem masculines treated there), distances it from the canonical case. Situations in which the distinct inflectional classes involved are each used in full to provide inflectional material for a larger paradigm (e.g. the regular Ancient Greek phenomenon whereby adjectives with a citation form in *-os* make use of second declension morphology in the masculine/neuter and first declension morphology in the feminine, cf. Sihler 1995: 348) can also be treated as much further removed from the conceptual core of heterocclisis than those in which, as with *domus* or the Gothic *i*-stem masculines, only some forms from each of the

inflectional classes surface in the heteroclite paradigm (cf. the notion of ‘partial competition’ at Stump 2006: 305).²⁹

Other dimensions of canonicity which can be ascribed to heteroclisis along the lines laid out by Corbett and others will, on occasion, be noted in the course of discussion of the attested instances of heteroclisis treated in this thesis. This work does not primarily seek to identify highly canonical examples of the phenomenon in Latin and the Romance languages: for example, as stated above, I avoid unique cases (as exemplified by *domus*) in favour of those which have a larger impact on the verb system in general, but which are thus less canonical along this particular dimension. Nonetheless, it is valuable to recognize that heteroclisis is amenable to a canonical treatment just like the other inflectional phenomena discussed in the foregoing for which this has been explicitly stated; what is more, this viewpoint can go some way towards reconciling the different approaches to the use of the term ‘heteroclisis’ itself shown by Noyer and Stump, to whose theoretical models I now turn.

²⁹ Cf. also the different use of terminology in Stewart and Stump 2007, mentioned below at fn. 37.

2.4 Noyer 2005 on heterocclisis and ‘systematic mixed inflection’

As was mentioned above, Noyer 2005 treats the topic of heterocclisis (as understood here) under two separate headings according to the extent to which it displays coherence of patterning within the inflectional system, and in fact Noyer spends little time on the less systematic type of behaviour for which he reserves the label ‘heterocclisis or metaplasms’.³⁰ This terminological question should not disguise the article’s relevance for the study of the phenomenon at issue in this thesis: the detailed generative model presented by Noyer aims to account for much of the same inflectional behaviour which is also treated in Stump 2006, though the substantial difference in emphasis between the two reflects their divergent assumptions about the nature of morphology. At the same time both scholars explicitly, but for different reasons, distance themselves from the No Blur Principle of Carstairs-McCarthy.

Noyer’s discussion of heterocclisis is couched in terms of Distributed Morphology (‘DM’; Halle and Marantz 1993, 1994), a Minimalism-based theory of the architecture of grammar which counts him among its main proponents (see e.g. Embick and Noyer 2007, a programmatic overview of the approach). The fundamental intuition underlying DM is that morphology and syntax share commonalities, such as hierarchical structure and the relevance of morphosyntactic features like case and tense, which are best explained by incorporating morphology into the existing Minimalist account of syntactic behaviour rather than granting it a separate component of its own. The word is thus not a theoretically significant unit of analysis: words, like phrases, are generated by the syntax, and there is no fundamental distinction between what in some other theories would be treated as the

³⁰ In fact Noyer is not entirely consistent in reserving ‘heterocclisis’ for sporadic behaviour: for example, on p. 297 he refers to ‘heterocclitic inflection’ in the whole class of Ancient Greek nouns he labels the masculine **es*-stems.

combination of morphemes to make words and the combination of words to make phrases.

Underlyingly, then, according to the DM view words have a syntactic internal structure and basically consist of concatenated elements associated with the terminal nodes of a syntactic tree. An abstract feature at a terminal node, such as [+plural], will be realized by a piece of phonological form ('Vocabulary Item', VI), such as /z, s, ɪz/: in this illustrative case from English, the VI is inherently specified as affixal, but that is a contingent fact about English which need not have been so. The theory is often described as 'piece-based', because it posits that lexical and functional morphemes are realized in a similar fashion by these originally discrete elements; Stump's (2001: 1) label 'lexical' for DM in his typology of morphological theories refers to the same characteristic, on the grounds that all such 'pieces' are treated equally as items in a lexical list.

There is no reason to expect inflectional morphology to have any special status on such an account. However, as we have seen above, inflected words may empirically fail to show an internal structure reflecting the full complex of features which the syntactic system is apparently calling upon them to realize, or may show purely formal distinctions which reflect nothing in the syntax at all – neither of which would be predicted by the simplest version of a concatenative model. The DM model therefore provides for various properties and mechanisms which can generate the non-canonical behaviour so commonly observed in inflection, some of which are relevant here. For one thing, the syntactic principles envisaged by DM may lead a terminal node to bear several grammatical features at once, which can then all be realized by a single VI, for example English -s for [+3 +singular +present]. The framework also permits the realization of features by means of zero affixes, i.e.

elements with no phonological content at all, while non-affixal inflectional marking is treated as a type of stem allomorphy (which, by hypothesis, cannot itself realize grammatical features in DM) effected by ‘readjustment rules’: often both zero affixes and readjustment rules are required together, to account for alternations such as English *foot* – *feet*- \emptyset alongside regular concatenative *hand* – *hand-s*, where in DM terms the same feature [+plural] is being realized affixally in both cases.

DM also allows for the existence of inflectional classes, embodied by ‘diacritic’ (arbitrary) class features borne by particular stems which serve to determine the affixes those stems may combine with.³¹ In English, for example, a possible DM analysis might identify a class of verbs characterized by the diacritic feature [strong], such as *beat*, *hit*, *run*, which take a zero affix in the past tense as opposed to the *-d* selected by regular verbs (abstracting away from allophony in the regular affix). The difference in behaviour between strong verbs and others would be captured by positing the following VIs:

(2.6) $-\emptyset \leftrightarrow [+past] / [strong] + _$

(2.7) $-d \leftrightarrow [+past]$

The first of these says that the affix \emptyset realizes the terminal morpheme [+past] only in the context of a stem bearing the diacritic feature [strong]. Meanwhile, *-d* is simply listed as realizing [+past], with no further specification. It is a principle of DM that where several VIs exist which could realize the same terminal morpheme, the grammar will favour the most narrowly defined item which is still compatible with the complex of features to be realized: thus where a past-tense affix is required for a

³¹ It might seem that any grammatical theory aiming for ‘descriptive adequacy’ must make use of some notion corresponding to inflectional class, but for a counterexample see Emonds and Spaelti 2005, which denies that Latin possesses declensions in the noun, in accordance with an outlook that ultimately wishes to do away with the idea of inflectional class features altogether. This is the logical end-point of a syntactic approach to morphology, but even in the test case of Latin nouns this view requires phonological complications which evidently recapitulate in a thinly disguised form generalizations which are better treated as morphological.

strong verb such as *beat*, the zero affix will be selected ahead of *-d*, while only the latter, representing the default or elsewhere case, is available for verbs which do not bear the inflectional class feature [strong]. Defaults and underspecification in fact play an important role throughout DM, and are used to explain some types of systematic syncretism: for example, the fact that English present tense forms which are *not* 3sg all bear the same ‘zero affix’ is naturally captured by the following pair of rules, where any person-number combination other than 3sg is compatible only with the underspecified item in (2.9).

(2.8) $-s \leftrightarrow [+3 \text{ +singular} \text{ +present}]$

(2.9) $-\emptyset \leftrightarrow [+present]$

The final mechanism to be introduced here, impoverishment, also provides an account for certain systematic syncretisms. Impoverishment is the deletion of a morphosyntactic feature in a particular context, meaning that it plays no part in the selection of the appropriate VI to realize the underlying complex of features: its effect is to render too narrowly defined any VI specified for that feature. Bobaljik 2002 uses impoverishment to explain the syncretism between 2sg and 3sg in the past tense of Macedonian verbs: according to this account, the feature [+2] is deleted in the presence of the feature [+past], as laid out in the rule given in (2.10), with the result that the specifically second person ending *-š* found elsewhere in Macedonian verbal inflection is inapplicable and the default inflection shown by the 3sg form is selected instead.

(2.10) $[+2] \rightarrow \emptyset / [+singular] + [+past] + _$

The purpose of Noyer’s article is to demonstrate how the phenomenon of heteroclisis, in the broad sense, can be accommodated in DM in terms of these existing formal mechanisms, already attributed to the grammatical system in order to account for other morphological phenomena. As will be seen, the upshot of Noyer’s

treatment is that some heteroclite behaviour can be identified as nothing other than a particular type of *syncretism*, arising from the same principled characteristics of the model which underlie syncretism as it is more generally recognized. Those genuinely isolated instances which constitute Noyer’s ‘heterocclisis’, however, are to be taken purely as the result of listing. Noyer gives the example of Ancient (Attic) Greek *déndron* ‘tree’, whose inflected forms nearly all pattern according to the thematic declension, but whose usual dative plural *déndressi* shows an affix *-si* which distinctively characterizes the athematic declension.³² This fact is captured in DM simply by stating the stem *dendres-* disjunctively alongside the athematic declension among the contexts in which the affix *-si* applies, as in (2.11), where [-genitive -direct] is the featural combination used by Noyer, for unrelated reasons, to characterize the dative case.

(2.11) *-si* ↔ [-genitive -direct +plural] / [-thematic] OR [*dendres-*, ...] + _

It is striking that this formalization effectively expresses the unexpected appearance of the form *déndressi* as a fact about the ending *-si*, rather than a fact about the word *déndron* or its stem. The disadvantage of this viewpoint is clearer when several inflected forms of a word require the same treatment. Supposing, for the sake of argument, that Latin *domus* can be taken as inherently belonging to the second declension, a DM account would have to specify the stem *dom-* (or the like) among the listed contexts of application of the VIs responsible for several fourth declension endings – dative singular *-uī*, nominative plural *-ūs*, and so on – without any way to acknowledge that the presence of listed *dom-* in their entry is something they all share, producing a consistent link between *dom-* and fourth declension behaviour.

This drawback seems to follow naturally from the architecture of DM, in which a

³² I simplify Ancient Greek nominal inflection, and Noyer’s treatment of it, substantially here as the full details are extensive and do not affect my presentation. As Noyer recognizes, the stem extension *-es-* seen on *dendres-* itself requires a DM account.

lexeme's 'inflectional paradigm' has the same (lack of) theoretical status as a collection of related syntactic phrases.

Alongside simple listing for these idiosyncratic cases, meanwhile, Noyer details three separate ways in which the machinery made available by the DM model can account in a principled fashion for more systematic instances of heterocclisis (Noyer's 'systematic mixed inflection') shared by coherent groups of lexemes and, as a result, predict which patterns of heterocclisis can be generated by the linguistic system and which are ruled out. His schematic illustration of systematic mixed inflection is given in Table 2.12, where the italicized letters represent affixes (note that his labels 'case 1', 'case 2', 'case 3' assume that the lexemes involved are nouns, as are all the concrete examples he discusses):

Table 2.12 Schema for Systematic Mixed Inflection (modified slightly from Noyer 2005: 274)

	class I	class I 'special'	class II	class III
case 1	<i>a</i>	<i>a</i>	<i>w</i>	<i>w</i>
case 2	<i>b</i>	<i>b</i>	<i>x</i>	<i>v</i>
case 3	<i>c</i>	<i>y</i>	<i>y</i>	<i>y</i>

Of interest to us are those lexemes which Noyer labels as belonging to a 'special' subtype of inflectional class I in this schema, on the assumption that in this scenario they are numerous enough for their behaviour to qualify as generated by rule rather than listing. These do display the expected class I affixes *a* and *b* for cases 1 and 2 respectively, but in case 3 they take an affix *y* normally associated only with other classes. What requires explanation in the simple situation presented here, then, is just the selection of affix *y*.

In some instances of this type, Noyer points out, the selection of *y* rather than *c* in case 3 for a subset of lexemes in class I may be predictable from the phonological context. Phonologically conditioned alternation of this kind is a

commonplace phenomenon presumably acknowledged by all morphological models, and may involve two affixes which themselves are entirely dissimilar in shape (Carstairs 1988b). Nonetheless, in principle the existence of automatic effects such as these does not compromise the *morphological* unity of the inflectional class involved. Hence Noyer claims that in the ‘special’ version of class I in this schema, as long as phonological conditioning can be held responsible for the selection of *y* in case 3, ‘we have no mixed inflection class in reality’ (p. 276). It is certainly true that for an efficient formal analysis of a pattern like that seen in Table 2.12, if the phonological environment fully accounted for the alternation between *c* and *y* in class I then it would be superfluous to invoke the morphological notion of heteroclisis as well. However, such an analysis would in effect claim that speakers of the language involved made no connection between the affix *y* found on some lexemes of class I, to be taken merely as a contextual variant of *c*, and the identical affix *y* found on all lexemes of classes II and III: the analysis would be the same if *c* instead had some unique variant form *z* under the ‘special’ phonological conditions. This seems psychologically unrealistic, and in particular cases there may be good evidence to show that speakers do make the morphological connection here: precisely this situation is found in the case of the Romanian verb *a țese*, and others, treated in chapter 4. It should be pointed out, however, that this problem should not be attributed to any feature particular to DM or Noyer’s presentation itself: inevitably, such ‘hard cases’ where formal economy of analysis disagrees with psychological plausibility will be faced by any linguistic theory which distinguishes between phonological and morphological effects but allows for phonological conditioning on morphology.

In the cases covered by Noyer's first mechanism, phonological conditioning, the appearance of heteroclisis is essentially treated as 'accidental' in synchronic terms, as the affix *y* is simply a contextually restricted element belonging to class I that happens to look like an element belonging to another class or classes. The second mechanism he invokes, 'default spell-out' (i.e. default realization), also effectively allows for the impression that certain lexemes show incompatible inflectional class behaviour over the different forms they take, without requiring any such incompatible behaviour in the underlying features generating those forms. In the relevant instances, affix *y* is the *default* realization of case 3, applied when no more narrowly defined affix exists which is compatible with the features borne by the stem. That is, while the affixes *a* and *b* have entries (2.12) and (2.13) in the grammar, explicitly restricting their use to lexemes in class I, affix *y* has entry (2.14):

- (2.12) $a \leftrightarrow [\text{case 1}] / [\text{class I}] + _$
 (2.13) $b \leftrightarrow [\text{case 2}] / [\text{class I}] + _$
 (2.14) $y \leftrightarrow [\text{case 3}] \quad (\textit{elsewhere})$

There is thus no disagreement in inflectional class between *a*, *b* on the one hand and *y* on the other. In this case it is the behaviour of the 'non-special' subclass which requires an explanation. This is provided by entry (2.15) for the affix *c*, which states that it is appropriate for class I lexemes bearing the further feature [-F] (supposing that [+F] characterizes the 'special' subclass).

- (2.15) $c \leftrightarrow [\text{case 3}] / [\text{class I } -\text{F}] + _$

However, as Noyer points out, the slightest of variations on the inflectional schema illustrated above can give rise to patterns which cannot be accounted for in this way. In Table 2.13, for example, case 3 is realized in class III not by the affix *y* but by *c*, as it is in one subgroup of lexemes from class I. What this means in Noyer's terms is

that it is no longer possible to treat *y* as the default affix for case 3,³³ with the result that an explanation in terms of default realization cannot apply.

Table 2.13 Second schema for Systematic Mixed Inflection (modified slightly from Noyer 2005: 277) [compare Table 2.12 above]

	class I	class I 'special'	class II	class III
case 1	<i>a</i>	<i>a</i>	<i>w</i>	<i>w</i>
case 2	<i>b</i>	<i>b</i>	<i>x</i>	<i>v</i>
case 3	<i>c</i>	<i>y</i>	<i>y</i>	<i>c</i>

Attested patterns of this kind, Noyer claims, thus require a third type of explanation, which invokes the mechanism of impoverishment described above, as well as a different kind of default. In such cases, the 'special' lexemes of class I, which unexpectedly take the ending *y* in case 3, must possess a feature [+F] which is identifiable on independent grounds (in all the examples presented by Noyer, this is a gender feature). What is more, class II in this schema must be the *default class* for lexemes bearing this feature, again a fact which can be determined independently on the basis of other characteristics of the language. This is captured in DM by a redundancy rule, which targets any [+F] stem unspecified for inflectional class and transfers it into class II. The heteroclit behaviour of a 'special' class I lexeme is thus generated as follows. First, the inflectional class feature labelling the lexeme's stem as belonging to class I is deleted by impoverishment in the context of case 3, as shown in rule (2.16).³⁴ Subsequently, the resulting impoverished stem is subject to

³³ In fact this is not strictly accurate: it would be possible to retain *y* as the default simply by specifying that *c* is available in two disjunctive conditions, namely when realizing case 3 either on a class I or a class III stem. It is not clear whether Noyer wishes to rule out disjunctive listing in principle, as he suggests it as a possibility on p. 283, in reference to the distribution of instrumental singular endings in Russian. The DM analyses of Latin and Armenian nominal inflection in Halle and Vaux 1998 also make liberal use of disjunctive listing. Clearly, though, allowing for disjunctive listing in the conditions on Vocabulary Items would sacrifice any predictive power aimed for by Noyer's approach.

³⁴ Strangely, Noyer's formalization reproduced in (2.16) suggests that the feature [+F] is also deleted by impoverishment, but evidently this is unintended, as the feature must 'survive' in order to trigger the redundancy rule in (2.17). I assume that this is merely an oversight in Noyer's use of the notation, and does not reflect any theoretical obstacle in DM to the type of impoverishment required.

the redundancy rule which determines that class II is the default for [+F] stems, as shown in (2.17):

(2.16) [class I +F] → Ø / _ + [case 3]

(2.17) [+F] → [class II]

This derivational sequence will give rise to the selection of the class II ending *y* on the ‘special’ lexemes of class I.

Naturally, this third mechanism can be invoked even in cases where the previous mechanism of default spell-out is not logically ruled out. A concrete instance explained by Noyer in terms of impoverishment concerns the inflectional behaviour of a class of masculine nouns in Gothic, labelled the *i*-stem masculines (a label which anticipates the eventual analysis). As shown in Table 2.14, these nouns show no distinctive morphology of their own, but in different parts of the paradigm they make use of inflectional endings proper to two distinct classes which must be posited independently, namely the *i*-stem feminines and *o*-stem masculines: all plural forms pattern with the former, while the genitive and dative singular clearly side with the latter (the nominative and accusative singular endings are ambiguous). In line with the previous mechanism, the explanation could be given that the apparent *i*-stem endings (*-eis -ins -ē -im*) are in fact simply the default nominal endings in the plural; or alternatively that the apparent *o*-stem endings (*-s -Ø -is -a*) are simply the default in the singular, but either gambit would lack independent motivation. Noyer points out, however, that the *o*-stem class as a whole can be treated as the default class for masculines, on the grounds that it provides the usual endings signalling masculine agreement on the adjective (which Noyer takes to be inherently unspecified for inflectional class in Gothic). On this basis, the group of nouns in question here can be taken as the ‘special’ [+masc] subset of a single *i*-stem class, the equivalent of Class I in the abstract schemas above, with *o*-stem behaviour in the singular resulting

from impoverishment of this inflectional class feature followed by the application of the redundancy rule transferring masculines into the *o*-stem class. The required impoverishment and redundancy rules (2.18) and (2.19) correspond to the more abstract (2.16) and (2.17) respectively:

(2.18) [*i*-stem +masc] → Ø / _ + [+sing]

(2.19) [+masc] → [*o*-stem]

Table 2.14 Mixed inflection in Gothic masculine *i*-stem nouns (modified from Noyer 2005: 275)

	<i>i</i> -stem (f)	<i>i</i> -stem (m)	<i>o</i> -stem (m)	<i>u</i> -stem (m/f)
nom sg	-s	-s	-s	-us
acc sg	-Ø	-Ø	-Ø	-u
gen sg	-áis	-is	-is	-áus
dat.sg	-ái	-a	-a	-áu
nom pl	-eis	-eis	-ōs	-jus
acc pl	-ins	-ins	-ans	-uns
gen pl	-ē	-ē	-ē	-iwē
dat pl	-im	-im	-am	-um

Within the confines of the DM model, the appearance of heteroclisis in a coherent group of lexemes can come only as the result of one of these three mechanisms, of which only the second and third can be responsible for genuinely morphological (rather than morphophonological) effects. That is, default spell-out and impoverishment, already central features of DM serving to account for patterns of syncretism within the paradigm, are also responsible for what can be seen as ‘syncretism’ *between inflectional classes* in particular cells of the paradigm. The DM approach thus predicts that only certain types of ‘interclass syncretism’ should be found, owing to the way in which these mechanisms operate. As we have seen, the inflectional affix which marks out a group of lexemes as heteroclite (i.e. affix *y* when found on forms assigned to the ‘special’ subset of class I, in the schemas provided here) must either be the overall *default affix*, or the affix of the *default class* assigned

to stems bearing the same feature [+F] as those in the heteroclite group. This is Noyer's *Interclass Syncretism Constraint*.³⁵

Noyer himself notes (p. 304) the resemblance between this constraint and the earlier No Blur Principle of Carstairs-McCarthy, which likewise proposes restrictions on the lexical distribution of rival inflectional affixes. Carstairs-McCarthy's proposal clearly allows for fewer expected patterns of behaviour than Noyer's, as the former does not invoke the notion of default inflectional classes alongside default affixes. But as Noyer points out, this does not make No Blur the more restrictive theory. As No Blur is envisaged as the result of general tendencies in acquisition, it does not rule out the existence of counterexamples altogether, but merely entails that they should represent marked and unstable states of affairs. It is thus an open question to what extent the counterexamples to No Blur which have been adduced in the literature should be taken to vitiate the principle: as we have seen above, its status is a matter of debate. Noyer himself does not rule out some role for the No Blur Principle, although he himself provides a counterexample (p. 305) involving the instrumental plural endings of Old Russian.

By contrast, in terms of the DM model presented in this section, to say that a pattern does not obey the Interclass Syncretism Constraint is to say that there is no way for it to be generated by the grammatical system: such patterns simply cannot exist. Any systematic behaviour which violates this constraint is thus to be taken seriously as a threat to the model presented. Nonetheless, we will see in section 4.4 below that a well-attested instance of 'interclass syncretism', in Noyer's terms, in the

³⁵ The precise wording of this constraint is as follows (Noyer 2005: 278): 'Let *A* and *B* be distinct inflectional classes, where the morphosyntactic feature [\pm F] partitions the set of stems in *A* into two subclasses. If only the [+F] stems in class *A* share an affix α with the stems in class *B*, then either: (i) α is a default affix, or (ii) *B* is the default inflectional class for [+F] stems.'

standard Romanian verb system cannot be modelled by means of the DM treatment seen here. This failure appears to undermine the validity of the DM approach.

2.5 Heterocclisis in Paradigm Function Morphology (Stump 2006)

The general approach towards inflectional morphology adopted by Stump is very different from that espoused in DM, and accordingly Stump's treatment of heterocclisis is also strikingly unlike what we have just encountered. The fundamental difference between the two views, as is articulated by Stump himself (2001: 1-3) in a taxonomy of morphological theories which has often been drawn on by morphologists of various stripes (e.g. Müller, Gunkel and Zifonun 2004: 11; Hinzelin and Gaglia 2012: 16f.), is that while DM is a *lexical* theory, treating inflectional morphemes effectively as items from a lexical list which require direct realization in their own right, Stump's own theory of Paradigm Function Morphology (PFM) is *inferential*, meaning that the form taken by an inflected word results from the application of *rules* to a lexical root which together realize all the morphosyntactic features it bears. The upshot of this is that in PFM 'morphemes' do not exist as independent syntactic elements or distinct nodes in the syntactic tree, and the syntax is not sensitive to the internal morphological structure of words; instead, words are 'syntactic atoms' all of whose morphosyntactic properties are together associated with a single node in the syntax. It is thus a theory of 'a-morphous' morphology in the sense of Anderson 1992.

In order to account only for the simplest cases, one could imagine an inflectional system involving such rules as operating in the following way: a root (default stem), specified for a set of morphosyntactic properties, would be subject to rules of stem formation and realization applying in sequence. For example,³⁶ to produce the nominative singular *amīcus* of the Latin masculine noun meaning 'friend', the morphology would apply to the pairing of root and morphosyntactic

³⁶ This illustration is based on Stewart and Stump (2007: 392).

property set ⟨amīc, [masc nom sg]⟩ the following rules, just two out of the many morphological rules relevant to the full inflectional paradigm of this noun:

a. Stem-formation rule

Where root R is a second declension nominal, R's thematized stem is *Ru*.

b. Realization rule

Where X is the thematized stem of a second declension root R, and R is an adjective or masculine noun, the cell ⟨R, {masc nom sg}⟩ is realized as *Xs*.

On the basis of the pairing ⟨amīc, {masc voc sg}⟩, meanwhile, a different realization rule would produce the vocative singular *amīce*, based directly on the root *amīc* rather than the thematized stem. The full inflectional paradigm of *amīcus* would be produced in this way by general rules applying to a specified portion of the lexicon.

The cases of *amīcus* and *amīce* are straightforward because, for each of them, there is no difference between the set of morphosyntactic properties relevant to its syntactic behaviour and the set relevant to its morphological realization. Such a state of affairs would be universally characteristic of the *canonical* inflectional system, in the sense used above. However, in PFM, certain types of non-canonical phenomenon are modelled by allowing a theoretical distinction to be drawn between the two sets of properties. For example, neuter nouns in Latin, as widely across Indo-European, never show any inflectional distinction between their nominative and accusative forms, instead showing a syncretism which runs systematically through the language (e.g. *dōnum* 'gift, nom/acc sg', *dōna* 'gifts, nom/acc pl'). On the other hand, the distinction between nominative and accusative evidently remains relevant to Latin syntax, which does not alter in its operation according to whether a given noun's nominative and accusative forms are realized identically or not. A model of the Latin inflectional system must therefore capture the fact that it only ever provides a single form catering to any two pairings of roots and morphosyntactic property sets which differ only in that one contains {neuter nom} while the other contains {neuter acc}.

This is achieved in PFM by the distinction of *content paradigms* from *form paradigms* – the first providing an interface with the syntax, associating a *lexeme* with its morphosyntactic properties relevant to syntactic behaviour, and the second providing an interface with the realizational morphology, associating a *root* with the morphosyntactic properties relevant to inflectional form. As a default, the morphosyntactic properties involved will be identical in the two types of paradigm. But in the current case, while the content paradigm of *dōnum* contains distinct nominative and accusative cells, the form paradigm does not. Cells in the two types of paradigm thus appear as follows, showing how PFM expresses this syncretism by means of the OR operator \vee . It is the relevant cell in the form paradigm to which stem-formation and realization rules then apply, giving the nom/acc sg form *dōnum* itself.

Content paradigm of the lexeme DŌNUM ‘gift’:

a. $\langle \text{DŌNUM}, \{\text{neut nom sg}\} \rangle$

b. $\langle \text{DŌNUM}, \{\text{neut acc sg}\} \rangle$

...

Form paradigm of the root *dōn* ‘gift’:

a. $\langle \text{dōn}, \{\text{neut nom} \vee \text{acc sg}\} \rangle$

...

Syncretism in Latin neuter nouns is one kind of absence of one-to-one mapping between the form content and the content paradigm. This is captured in the special Latin rule of *paradigm linkage* given below. In contrast, the simplest cases (as illustrated by *amīcus*, *amīce*) show the simplest possible mapping between the two paradigms, which can be treated as the cross-linguistic default, applying if no more specific rule of paradigm linkage is applicable.

Latin rule of paradigm linkage

Where the lexeme L has R as its root, and α = sg or pl:

$\langle \text{L}, \{\text{neut nom } \alpha\} \rangle, \langle \text{L}, \{\text{neut acc } \alpha\} \rangle \Rightarrow \langle \text{R}, \{\text{neut nom} \vee \text{acc } \alpha\} \rangle$

Universal default rule of paradigm linkage (cf. Stump 2006: 286)

Where the lexeme L has R as its root, and σ is a morphosyntactic property set:

$$\langle L, \sigma \rangle \Rightarrow \langle R, \sigma \rangle$$

Paradigm linkage, illustrated here to show its use as a way to account for syncretism, is also used in the PFM model to account for deponency, such as voice-deponency in Latin as seen in section 2.1 above. This is captured straightforwardly in PFM by the following rule of paradigm linkage, which overrides the default (cf. Stewart and Stump 2007: 394):

Latin rule of paradigm linkage

Where L is a deponent verb having root R:

$$\langle L, \{\text{active...}\} \rangle \Rightarrow \langle R, \{\text{passive...}\} \rangle$$

Schematically, then, the paradigm linkage treatment of deponency in PFM involves a paradigm linkage of the form $\langle L, \sigma \rangle \Rightarrow \langle R, \sigma' \rangle$, in which σ' represents a different set of morphosyntactic properties from σ .

Both the above types of non-canonical inflectional behaviour are thus modelled as instances of the alteration of the morphological property set of the content paradigm in its link with the form paradigm. However, the way is equally left open for special deviations from the default as regards the *left-hand* side of the pairing, involving not the morphological property set but the root to which the appropriate rules are applied. The mechanism adopted by PFM to account for heteroclisys takes advantage of this. The crucial point is that stems may be specified for inflectional class: thus the root *amīc* of *amīcus* above was specified as belonging to the second declension, and was thus subject to certain particular stem formation and realization rules. In instances of heteroclisys, then, PFM sees the relevant rule of paradigm linkage as manipulating the stem found in the form paradigm in such a way that it bears different inflectional class features in different cells. To address a concrete example introduced above (cf. Table 1.2), Stump proposes (2006: 289) that

the lexeme *pramen* has a root *pramen*_{sm}, marked as belonging to the soft-masculine declension, which appears in the singular cells of the form paradigm, but also a different ('coradical') stem *pramen*_{hm}, belonging to the hard-masculine declension, whose appearance in the paradigm is brought about by a special rule of paradigm linkage which applies to nouns behaving like *pramen* and overrides the default:

Czech rule of paradigm linkage (based on Stump 2006: 289)

Where L is a nominal lexeme that belongs to the *pramen* class and has S as its hard-masculine coradical:

$\langle L, \{\text{plural...}\} \rangle \Rightarrow \langle S, \{\text{plural...}\} \rangle$

The presence of these distinct stems, formally identical except for their inflectional class affiliation, in different cells of the form paradigm of *pramen* is what underlies, in Stump's terms, the heteroclite behaviour this lexeme is observed to show.

Importantly, PFM makes use of the same approach to account for more systematic instances of heteroclisis (which would thus not fall under that heading in Noyer's terms). For example, Stump makes use of specialized rules of paradigm linkage to treat a set of Czech nouns behaving analogously to the *i*-stem masculines of Gothic (e.g. *předseda* 'president'), which Noyer counts as displaying 'systematic mixed inflection'; what is more, the 'recruitment' of distinctions of declension in order to distinguish between genders within the paradigm of the Ancient Greek adjective, also mentioned above (end of section 2.3), is likewise treated in terms of the selection of a 'feminine coradical' (belonging to a different declension from the masculine root) which is assigned to parts of the form paradigm characterized by the morphosyntactic property set {feminine...}.³⁷

³⁷ The availability of this general approach as a treatment of all these phenomena helps to motivate Stump's (2006) use of 'heteroclisis' (as noted above, section 2.1) as a blanket term for them all. However, it is notable that in Stewart and Stump 2007 cases such as the recruitment of declension classes for different parts of the adjectival paradigm are no longer discussed under the heading of heteroclisis, which is reserved for cases where the use of more than one inflectional pattern within the paradigm is not an expected feature of the type of lexeme in question.

Heteroclisis can thus be treated in the PFM model as the outcome of language-particular instantiations of a general mechanism, paradigm linkage, which is taken as required by the architecture of the grammar, and whose existence is already implicated in the explanation of other non-canonical inflectional characteristics such as syncretism and deponency. In particular, however, as Stump notes (2006: 283), this theoretical account of heteroclisis effectively treats it as a particular type of *suppletion*. This contrasts with the tack taken by the DM approach, in which (some) ‘systematic mixed inflection’ is viewed as a type of (interclass) *syncretism*.

It is interesting to note, in both cases, the possibility of accommodating heteroclisis by means of mechanisms already set up within the theory to account for other non-canonical phenomena. However, the treatment of heteroclisis in PFM as a variety of suppletion is connected with a valuable further feature of the approach which is absent from the DM account. The PFM model explicitly takes account of the widespread existence cross-linguistically of paradigmatic stem alternations which do not express morphosyntactic distinctions but whose distribution over the paradigm is purely morphologically conditioned. Stump (2001: ch. 6) discusses such phenomena particularly with regard to nominal stem alternation in Sanskrit, where, for example, some nouns and adjectives bear a distinctive ‘strong’ stem, which appears throughout the nominative, vocative and accusative (and nowhere else in the paradigm) *except* in the case of neuter singulars and duals and masculine accusative plurals, a zone which clearly does not constitute a natural morphosyntactic class. Stump notes, citing Corbett,³⁸ that morphological conditioning of this kind has also

³⁸ Stump (2006: 293) gives the reference ‘Corbett (2006)’, which is listed in the bibliography as referring to the then unpublished article which eventually surfaced as Corbett 2009b. However, this article, in its final form, barely addresses the point at issue. See, however, Corbett 2007a (particularly pp. 18-23), which may be the intended reference.

been found to articulate the distribution even of strongly suppletive forms over the inflectional paradigm (cf. the example of French *aller* presented shortly on p. 108). In the same vein, the PFM model makes available the possibility for heteroclisis to be directly sensitive to these morphological patterns: the morphological label ('the strong property set'), already required in the morphology of the language to single out the relevant paradigmatic zone as it is affected by stem alternation, is given as part of the rule of paradigm linkage which gives rise to heteroclisis in the observed case.

It is not easy to see how the possibility of purely morphological conditioning on heteroclisis could be made available as part of the DM approach, which as it stands does not allow for the morphological coherence of 'morphosyntactically unnatural' groups of paradigm cells in this way. In one sense this is by no means a defect, since the lack of this possibility makes for a more restrictive theory. It may be clear from the preceding overview that the PFM approach to heteroclisis offers little by way of theoretical constraints: as Stewart and Stump (2007: 401) themselves comment, '[t]he notion of paradigm linkage is quite powerful', it is not clear whether there are any restrictions on the size or nature of the classes to which a given paradigm linkage rule may apply, and to an extent PFM thus seems to represent a more overtly 'theoretical' recasting of the descriptive enterprise of canonical typology, in which attested morphological behaviour is viewed against the backdrop of a neat grid defined by the available morphosyntactic properties. Further allowing both morphological and morphosyntactic conditioning on heteroclisis removes another potential constraint.

However, the widespread existence in natural language of paradigmatic zones whose cells pattern together morphologically in some respect, despite not being

unified by any functional or other extramorphological factor, is by now well documented. Alongside Stump's observations on the nature of stem distribution in Sanskrit and elsewhere, such purely morphological ('morphomic': Aronoff 1994) patterns have been identified e.g. for the verb systems of Sora (Stump 2005a) and Nepali (Bonami and Boyé 2010) and are ubiquitous across the morphology of Kayardild (Round 2012). Such morphomic coherence can equally be viewed in terms of the paradigm structure conditions of Natural Morphology: while reliable paradigmatic patterns often do in fact characterize morphosyntactically coherent groups, PSCs are not bound by principled restrictions on which cells of the paradigm may appear together in the same structural 'domain', meaning that there is nothing standing in the way of the development of purely morphological structure of this kind.

Notably in the context of this study, the emergence of such patterns in the development of Romance has been the subject of a great deal of research (cf. for recent examples Maiden et al. 2011; Cruschina, Maiden and Smith 2013). I mention two of these patterns here, which will feature in later chapters. In the Romance 'PYTA pattern' (the acronym stands for the Spanish term *perfecto y tiempos afines*, Maiden 2001), verb forms continuing the Latin perfect stem (cf. section 3.1) often continue to share a distinctive stem absent from the rest of the paradigm, even though they generally no longer possess any common function or morphosyntactic property: cf. the situation in Spanish, where preterite *hizo*, past subjunctives *hiciera/hiciese*, and the archaic future subjunctive *hiciera* pattern together against e.g. infinitive *hacer* (reflecting the opposition between Latin *fēc-* and *fac-*), on what are now purely morphological grounds. Meanwhile, in the Romance so-called 'N-pattern' (Maiden 2005), the verbs involved typically possess a special stem

restricted to the singular and 3pl of the present indicative and subjunctive (with the 2sg imperative), e.g. Spanish 1sg present indicative *muéstr-o* ('N-pattern') vs. 1pl *mostr-ámos*: this pattern arose as a result of the inherited difference in the position of stress on these forms (as marked), which gave rise to differential sound change in the vowel of the stem. Significantly, these patterns are not found only as 'fossils' continuing pre-existing distinctions but appear more widely in the languages involved, having evidently served as models for morphological innovation. Thus Corbett (2007a: 19) notes, following Aski 1995 and others, that the N-pattern determines the distribution of the two suppletive stems over the present indicative of French *aller* (*vais, vas, va, allons, allez, vont*), although both the etyma involved, *vadō -ere* and *ambulō -āre*, had full paradigms in Latin. We will see in later chapters that these morphomic patterns can be relevant to the distribution of Romance heteroclisis as well.

In light of this, it seems appropriate that PFM should provide a way to model the existence of purely morphologically defined groups of cells, and their potential role in determining the behaviour of heteroclisis or suppletion more generally over the paradigm. Indeed, on the basis of evidence to be discussed primarily in Chapter 3 of the present work, Maiden 2009 has surmised that, if one out of morphological and morphosyntactic conditioning is superfluous to the distribution of heteroclisis, it is the latter: that is to say that heteroclisis, where it is found, is only *ever* distributed along lines laid down by pre-existing patterns of stem alternation, and can only appear to be sensitive to morphosyntactic distinctions where stem alternation already follows those lines.

The investigation of this question necessarily involves a *diachronic* treatment of the issues involved. The influence of inherited morphological patterns on the

wider inflectional system can only manifest itself by means of diachronic morphological change – as is implicit, for example, in the example of the paradigm of French *aller*, where the synchronic suppletion observed comes about as the result of the historical process which can be called (with Juge 1999, who treats analogous issues in Ibero-Romance) *incursion* of forms into the paradigm along N-pattern lines. On the other hand, it has been observed that suppletion can arise in a paradigm in the absence of a historical process of incursion, as the result of differential sound change (Plank's (1996) dissimilatory, as opposed to combinatory, suppletion): this would not necessarily be expected to proceed along lines previously laid down by existing morphomic patterns.

This diachronic angle is absent from Stump's presentation of heteroclisis. However, it is clear that diachronic investigation of questions concerning the distribution, but also simply the origin, of different cases of heteroclisis has the potential to aid our understanding of the synchronic phenomenon, as well as addressing a topic of interest for the history of the individual languages concerned. And it is a sign of the growing scholarly recognition of the modern, typologically useful conception of heteroclisis, as a phenomenon connected with inflectional class membership, that historical studies on individual languages are now starting to make explicit use of it; the present thesis stands as a part of this recent trend.

Before turning to the cases of Latin and Romance heteroclisis which feature here, I therefore present three recent diachronically-oriented studies which invoke heteroclisis in its modern sense to describe the behaviour of particular varieties.³⁹ The first of these, Karatsareas 2011a, details the rise of heteroclisis as a distinctive characteristic of the nominal system of Asia Minor Greek. The second, Luís 2008, in

³⁹ Maiden's proposal, mentioned just above, itself emerges from the historical study of heteroclisis in dialectal Romanian, but I leave this to one side for now, as its findings will be treated at length in Chapter 4.

intending to identify heteroclisis in the verb system of Caló (Spanish Romani), treats a case whose diachronic origins lie in Romance; however, as mentioned above, the application of the fundamentally synchronic notion of heteroclisis to diachronic questions leaves some room for confusion with regard to the claims which are being made, and Luís's study, which can in fact be seen to focus on a phenomenon somewhat different from heteroclisis, illustrates the potential of this issue to mislead. The final study to be discussed in this chapter, Esher 2012, is not devoted to heteroclisis in particular, but provides evidence for a particular phenomenon which illustrates the relevance of the notion of heteroclisis to the study of the Romance verb; I will point out that in broad outline, its findings can be compared to those seen in the case treated by Karatsareas.

2.6 Three recent studies invoking heterocclisis

2.6.1 Karatsareas 2011a on Asia Minor Greek

Karatsareas 2011a is a study of the origins and development of a notable characteristic of nominal inflection in Asia Minor Greek (AMGr), a label comprising the dialects of Modern Greek which were spoken in several parts of eastern Asia Minor up to the population exchange that took place between Greece and Turkey in 1923 (and still survive to varying extents in Greece or their original locations). These dialects show several inflectional classes in the noun, which can largely be identified historically with corresponding classes inherited by the other Modern Greek varieties. Accordingly, for our purposes here, the ‘starting point’ for the morphological development in question can be illustrated by means of representative paradigms from Standard Modern Greek (SMGr). Table 2.15 gives examples of three different declensions in the standard language, each of which is associated with nouns of a particular gender; however, for each gender there are also other declensions available in Greek, which are not presented here.

Table 2.15 Three nominal paradigms belonging to different declensions in Standard Modern Greek

	<i>peđí</i> (n) ‘child’		<i>ánthropos</i> (m) ‘man’		<i>θάλασα</i> (f) ‘sea’	
	sg	pl	sg	pl	sg	pl
nom	<i>peđí</i>	<i>peđjá</i>	<i>ánthropos</i>	<i>ánthropi</i>	<i>θάλασα</i>	<i>θάλασες</i>
voc	<i>peđí</i>	<i>peđjá</i>	<i>ánthrope</i>	<i>ánthropi</i>	<i>θάλασα</i>	<i>θάλασες</i>
acc	<i>peđí</i>	<i>peđjá</i>	<i>ánthropo</i>	<i>ánthropus</i>	<i>θάλασα</i>	<i>θάλασες</i>
gen	<i>peđjú</i>	<i>peđjón</i>	<i>ánthropu</i>	<i>ánthropon</i>	<i>θάλασας</i>	<i>thalasón</i>

The phenomenon under discussion centrally involves what can be called the neuter *i*-declension, exemplified here by *peđí* ‘child’. This declension largely comprises nouns whose stems were originally formed by means of the diminutive suffix *-i-* (e.g. *peđí* < Ancient Greek *paid-i-on*, cf. *paid-* ‘child’), which naturally characterized all inflected forms of the relevant lexemes, and accordingly in the majority of the Greek

dialects membership of the *i*-declension is a straightforward categorical matter. However, in the AMGr dialects, the situation is more complicated: as Karatsareas reports, it is common to find nominal forms inflected for genitive singular or genitive plural which clearly pattern with this *i*-declension, while other inflected forms of the noun involved are equally unambiguous in patterning with another of the existing declensions. The nouns involved may well be neuter themselves, but many masculine and some feminine nouns are also found to show this behaviour. Karatsareas cites an example of the type from the Cappadocian Greek variety of Axó (Mavrochalyvidis and Kesisoglou 1960: 200), where *miljú*, rather than expected *milu*, is found as the genitive singular of masculine *milos* ‘mill’.

Drawing on the presentation of the concept in Noyer 2005⁴⁰ and Stump 2006, Karatsareas identifies AMGr nouns of this kind as showing heteroclisis. The diachronic basis of this synchronic state of affairs is clear. Historical and comparative evidence confirms what could have been suspected purely on logical grounds, namely that it is the *i*-declension affiliation (common to all the nouns involved) that represents the innovation here: the characteristic morphology of the *i*-declension has been taken on in certain parts of the paradigm by a selection of nouns, such as *milos*, which have no connection with that declension etymologically. Heteroclisis in this case has thus been brought about by a process which can be considered its diachronic counterpart, namely incursion, or the adoption of morphology proper to one inflection class in just part of the paradigm of certain lexemes which originally belonged to another.

But several further diachronic conclusions about the AMGr dialects can be drawn on the basis of these heteroclite forms. For one thing, in analyses of the Greek

⁴⁰ Cited by Karatsareas as Noyer (2004).

nominal system, the characteristic *i* (alternating with *j* in SMGr as in most Greek varieties) of the *i*-declension nouns is generally taken as a (morphologically inert) part of the lexical stem, to which the appropriate inflectional endings are added (i.e. nominative/vocative/accusative singular *peði-Ø*, genitive singular *peđj-ú* etc.); and historically that was indeed the case. But as Karatsareas argues (against e.g. Ralli 2006: 136-41), the development seen in these heteroclite nouns suggests that, at least in the dialects at issue here, it has been reanalysed as part of the inflectional endings of the *i*-declension: this process is what makes available as such the endings *-jú*, *-jón* (or their dialectal equivalents) which are adopted in these cases of heteroclitism to give innovative forms such as *miljú* above.⁴¹

Meanwhile, comparison of the AMGr facts with findings from other dialects sheds light both on the early history of this pattern of heteroclitism and on the position of AMGr itself among Greek varieties. Karatsareas points to descriptions of the Northern Greek dialects of Lesbos (e.g. Papadopoulos 1926: 57) and Samos (Zapheiriou 1995: 91f.) which mention the existence in these dialects of unexpected *i*-declension forms in the genitive singular and plural of nouns otherwise belonging to other declensions, exactly as we have seen for AMGr: for example, Lesbian *áθrupus* ‘man’ (morphologically equivalent to SMGr *ánthropos* in the table above) has *i*-declension genitive singular *aθrupjú* and genitive plural *aθrupjún*. This striking agreement in behaviour suggests that a close relationship obtained at some historical stage between these dialects and those of Asia Minor, which has not generally been given its due. But at the same time, the Lesbian and Samian versions of this phenomenon are much more restricted in their lexical spread than any seen in AMGr.

⁴¹ The full argument for treating *ij* as a part of the ending is more complicated than this, and cannot be covered here: the essential point is that treating it as part of a special stem type associated with *i*-declension behaviour would imply unprecedented and otherwise unnecessary types of stem allomorphy in certain heteroclite nouns whose behaviour is accounted for straightforwardly on the reanalysis view, such as Oenoe Pontic *paxčás* ‘garden’, genitive *paxčadıu* (Karatsareas 2011b: 117f.).

In the Lesbian dialect the lexemes involved are drawn almost exclusively from a pool of masculine and neuter nouns belonging to what can be called the *os*-declension (*áθrupus* = SMGr *ánθropos*) and the closely related *o*-declension (*próvatu* ‘sheep’ = SMGr *próvato*) respectively; in Samian, only nouns from the neuter *o*-declension figure. What is more, those nouns which exhibit heteroclisis overwhelmingly share the same ‘proparoxytone’ stress pattern, in which stress falls on the antepenultimate syllable in the nominative, a constraint which clearly does not apply to heteroclite AMGr nouns such as *mílos*. The natural conclusion is that Lesbian and especially Samian conserve the earliest distribution of *i*-declension heteroclisis over the lexicon, while in the AMGr dialects its application has been substantially extended to give the modern situation.

Karatsareas proposes that this initial association of heteroclisis with proparoxytone nouns reveals the original motivation behind the phenomenon in this case. Nouns of this shape were subject to an inherited principle of stress shift in the genitive singular and plural, reflecting an earlier moraic constraint prohibiting antepenultimate stress in forms which historically contained a long final vowel (Probert 2003: 33); however, this behaviour clashed with a more recent tendency towards fixed stress in the nominal paradigm, which thus favoured antepenultimate stress in these forms. This would have given rise to uncertainty in the appropriate position of stress – which, however, was not felt in the *i*-declension, where stress predictably fell on the characteristic *i* itself in both genitive singular and genitive plural at the relevant historical stage.⁴² It has been demonstrated (see e.g. Sims 2007) that the existence of similar competition between two accentual patterns available for the genitive plural in the main feminine declension, that of *θálasa* in Table 2.15, is

⁴² In some dialects, including SMGr and Cappadocian (but not Pontic), stress was then transferred to a following vowel, giving e.g. *peǰjú* from earlier *peǰíu*.

responsible for the defectiveness which targets the genitive plural of thousands of feminine nouns in other dialects of Greek, including the standard. In light of this, the adoption of heteroclisis in the present case can be understood as another morphological strategy, less drastic than defectiveness, towards the resolution of formal instability in particular cells: speakers avoided uncertainty in the genitive singular and plural of some proparoxytonic nouns belonging to the neuter *o*-declension by innovatively granting them the predictable behaviour shown by the neuter *i*-declension. After this initial trigger, heteroclisis then spread lexically to affect nouns of other declensions to a greater or lesser extent in the different dialects.

2.6.2 *Luís 2008 on Caló*

Luís makes use of the notion of heteroclisis in her description of the verb morphology of Caló (Spanish Romani), which forms part of a wider survey of the effects of language contact on Romance verb paradigms (Luís 2008). Caló is a mixed language which developed in Andalusia among bilingual speakers of Romani and Spanish, and predominantly makes use of lexical material taken from the former and grammatical material (including inflectional material) taken from the latter, e.g. *jal-ar* ‘to eat’, where *jal-* is Romani while *-ar* can be recognized as the Spanish infinitive ending (Boretzky and Igla 1994: 55). However, as verbs in Romani obviously do not belong inherently to any particular Spanish conjugation, this structural principle is not enough on its own to determine *which* of the Spanish inflectional endings will be used on verbs in Caló. According to Luís (after Bakker 2003), the answer is as follows: all verbs show first conjugation inflections in the majority of the paradigm, thus infinitive *-ar* but also third person present *-a* and imperfect *-aba*; but the preterite is exceptional, showing (at least) 1sg *-í* and 2sg *-iste*,

which are characteristic of the non-first conjugation verbs of Spanish (cf. first conjugation *-é, -aste*).

In her categorization of the diachronic processes which may affect verbal inflection in cases of linguistic contact, Luís presents this state of affairs in Caló as the result of the levelling of conjugation classes, distinct from the other logical possibilities of loss, retention and extension attested in other contact varieties (p. 142). While Spanish verbs fall into several conjugations, these distinctions are neutralized in Caló, but not thanks to the survival of only one conjugation's inflectional material: rather, morphological levelling has removed the incompatibility between inflections belonging to different conjugations in Spanish, which are now yoked together to make up the single pattern of behaviour available to verbs in Caló.

At the same time, in synchronic terms Luís describes Caló verbs as showing *heterocclisis* (p. 146). In using this term, she certainly intends it in the sense involving inflectional class affiliation, as found in recent typological and theoretical literature and relevant to this thesis: when introducing the concept of heterocclisis, she cites Baerman's (2007: 16) example of Latin *balneum* 'bath', which inflects according to different declensions in the singular and the plural, and mentions that one recent approach to the phenomenon is laid out in Stump 2006, while her characterization of heterocclisis as '[t]he phenomenon whereby lexical items take stems which belong to distinct inflection classes' (p. 146) draws on Stump's own.⁴³ What is more, she explicitly points out the relevance of heterocclisis to questions about the organization of inflectional paradigms, which also shows that she takes it to be a synchronic label. For Luís, who notes that heterocclisis is found widely in the world's languages, the fact that it also exists in Caló represents one case where the morphological behaviour

⁴³ 'The term HETEROCLISIS refers to the property of a lexeme whose inflectional paradigm contains forms built upon stems belonging to two or more distinct inflection classes' (Stump 2006: 279).

of a contact variety is *not* in any way distinctive, but simply displays a characteristic also found in many ‘non-contact’ varieties.

In fact, however, the conjugational facts of Caló as presented by Luís give no reason to treat the language’s verb paradigms as showing heteroclisis in the sense she intends, and no light is shed on a comparison of contact varieties with other linguistic varieties by invoking the concept here. Intuitively it is easy to appreciate the connection evidently made by Luís between the behaviour of the Caló verb and the instances of heteroclisis documented by the scholars she cites: it rests on the fact that a Caló paradigm containing, among other inflected forms, both an infinitive *jal-ar* (not **jal-er*, **jal-ir*) and a 1sg preterite *jal-í* (not **jal-é*) seems to bring together inflectional morphology which belongs apart, just like the paradigm of *domus* or *balneum*. But there is a crucial difference which distinguishes the Caló case: namely, that the affiliation of infinitive *-ar* and preterite *-í* with distinct inflectional classes is not a fact about Caló at all, but about Spanish. It is only from the point of view of the Spanish inflectional system that there is any incongruity in the coexistence of forms such as *jalar* and *jalí* in the same paradigm. Meanwhile, as Luís demonstrates, within the verb system of Caló itself there is no division of lexemes into different conjugations: the distinction between them has been levelled out, and the result is precisely the phenomenon under discussion here, where the same inflectional behaviour is shown by all verbs.

Given this absence of distinct inflectional classes, heteroclisis in synchronic terms is logically impossible, as Luís’s own definition of the phenomenon makes clear. Her identification of this inflectional behaviour as heteroclite must be based not on synchronic but on diachronic considerations, concerning the Spanish origins of the morphological material involved. However, these are irrelevant to an account

of the organization of the verb paradigm in Caló and to the conception of heteroclisis which she is ostensibly interested in. The misapprehension shown by Luís here illustrates the problems that can arise from equivocation between the synchronic and diachronic perspectives in the identification of heteroclisis.

2.6.3 Esher 2012 on Occitan

I close this chapter by looking at a case where the notion of heteroclisis has been more fittingly applied to the study of Romance verbal phenomena. As mentioned above, Esher 2012 is not a study focusing on heteroclisis in particular; rather, it is a large-scale investigation into the morphological behaviour of the future and conditional in the Occitan varieties. As part of this investigation, however, Esher reports on a widespread phenomenon, which she identifies as heteroclisis, affecting these parts of the paradigm; as she observes, the context in which this phenomenon arises points to the likely motivation for the diachronic change which has brought about the current state of affairs.

For the purposes of this overview of Esher's findings, the verb lexicon of Occitan can be taken to fall into three conjugations, easily distinguishable by means of their infinitive form: verbs belonging to the first have an infinitive in stressed *-ar*⁴⁴ (e.g. *parlar* 'speak' [par'la]) and verbs belonging to the second⁴⁵ have an infinitive in stressed *-ir* (e.g. *finir* [fi'ni] 'finish'). The third conjugation is characterized by a wider variety of infinitive forms: a minority have an infinitive in stressed *-er* (e.g. *poder* [pu'de] 'be able'), but for the most part third conjugation infinitives end either in unstressed *-re* (e.g. *batre* ['batre] 'beat') or, less commonly, in unstressed *-er* (e.g.

⁴⁴ I provide reference forms in the standard Lengadòc (Languedoc) orthography; in general, the import of the forms involved is clear without the need for a phonemic transcription.

⁴⁵ To anticipate forthcoming chapters, this conjugation conventionally labelled the 'second' in Occitan, a convention followed by Esher, corresponds to what is known as the fourth conjugation in Latin and elsewhere in this thesis. Meanwhile, what is treated as the Occitan 'third conjugation' here reflects both the second and third conjugations of Latin.

plànher ['plane] 'pity'). It is the last of these groups, comprising the verbs with infinitive in unstressed *-er* (Esher's IIIb class), which is of interest here. Over much of the Occitan area, this class is limited to verbs with just a few different stem shapes, chiefly those with a final palatal (Esher 2012: 160).

The crucial feature of the verb system within which this phenomenon can be detected is the synchronic relationship which regularly holds in Occitan between the future and conditional forms on one hand and the infinitive on the other. Although the transparency of this relationship has been compromised by regular sound change in the paradigm of some verbs, particularly those with infinitive in stressed *-er* (e.g. infinitive *poder* 'be able', future *poirai* and not **poderai*), across the majority of the lexicon there remains a straightforward morphological correspondence between these parts of the paradigm. Accordingly, future and conditional forms show differences between conjugations which reflect those seen in the formation of the corresponding infinitives. For example, the Occitan future formation *parlarai* 'I will speak' [parla'raj] belongs distinctively to the first conjugation (cf. infinitive *parlar*), and differs in detail from a second conjugation future such as *finirai* [fini'raj] (cf. infinitive *finir*). What is more, within the third conjugation itself, the distinction just noted between verbs with infinitives in *-re* and those with infinitives in unstressed *-er* is also displayed in the future and conditional forms: and indeed these are the only parts of the paradigm other than the infinitive to make the distinction between these two subtypes of the third conjugation.⁴⁶ Examples in standard Lengadòc (e.g. Sauzet and Ubaud 1995: 90, 126) are *batre*, *batrai*, *batriá* vs. *plànher*, *planherai*, *planheriá* for infinitive, future and conditional respectively. In verbs of the *plànher* subtype, the future and conditional evidently display a distinctive third conjugation

⁴⁶ In most Gascon varieties, where this distinction is absent in the infinitive and only unstressed *-er* is found regardless of stem shape (e.g. infinitive *bàter*), it is likewise absent in the future and conditional (*baterèi*, *baterí*: Romieu and Bianchi 2005: 302).

theme vowel *e* which parallels the *a* and *i* seen in corresponding first and second conjugation forms respectively. At the same time, the absence of any such theme vowel in verbs of the *batre* subtype is also characteristic of the third conjugation in particular. The conjugational affiliation of each future/conditional formation is therefore straightforward as long as this system is preserved intact, as it is in most regional standards.

However, as Esher demonstrates, a widespread tendency exists in the Occitan area for speakers to produce innovative future and conditional forms of some verbs which can serve to complicate the straightforward relationship just described. The verbs affected are specifically those belonging to the subtype which shows an infinitive in unstressed *-er*. In place of conservative future and conditional forms containing the third conjugation theme vowel *e* (*planherai* etc.), numerous dialect descriptions record the existence of forms which show that a shift has taken place in favour of the characteristic morphology of a different conjugation. Thus for the dialect of Quercy, Allanche (1941: 51) notes that the third conjugation infinitive *plànger* (= Languedoc *plànher*) is accompanied by the future and conditional forms *planjarai*, *planjariái*, which pattern with the regular *first* conjugation verb *laurar* ‘work’ (*laurarai*, *laurariái*). The same shift is described for various dialects in the Occitan area, including Périgord (Benoît and Lavaud 1932), Limousin (Ronjat 1937) and Rouergue (Alibèrt 1976). Meanwhile, another, even more commonly attested shift displayed by the future and conditional of these verbs is towards the morphology of the *second* conjugation, with its theme vowel *i*: again, this development is not restricted to one geographical area but characterizes local dialects in many regions, including Languedoc (Alibèrt 1976, Lignièrès 1951: 54ff.), Drôme and Auvergne (Ronjat 1937: 314f., 317).

Crucially, Esher notes (p. 178) that these developments do not form part of a more general change in the nature of conjugational identity in the various varieties involved. It is not the case, for example, that in Quercy the inherited distinctions between third and first conjugation morphology are being effaced wholesale, in which case Allanche's observation would not specifically concern the future and conditional forms of verbs in *-er* at all. On the contrary, the phenomenon at issue here, although widespread geographically, is highly 'localized' within the conjugational system. Only future and conditional forms are involved, while the remainder of the paradigm is not generally affected by the shift. The innovative forms seen in these parts of the paradigm all show the introduction of non-third conjugation morphology into verb forms which historically belong to the third conjugation, and never the reverse. What is more, only those third conjugation verbs are affected whose infinitive is in unstressed *-er*. The only 'vulnerable' verb forms susceptible to this shift are those which at one point took the form *-erai* etc., *-eria* etc. (or their analogues in the different dialects).

The fact that this general type of innovation, seen in the verb paradigm of many Occitan varieties, is linked so specifically to inherited forms of a particular shape means that it is tempting to look first for an explanation which at least begins in the phonology. And, regardless of their status as morphological markers of three separate inflectional classes in the varieties which happen to be in question here, it is not hard to imagine that a pre-tonic *e* such as is found in the relevant verb forms might yield *a* or *i* for purely phonological reasons. A development of this kind would not necessarily be taken as *morphologically* innovative at all: compare the historical raising of unstressed *a* to *e* before *r* in Florentine (and hence standard Italian), which had the fortuitous effect of effacing the overt conjugational distinction between first

and third conjugation in the future and conditional (e.g. from *parlare*, *perdere*: *parlerò* < *-arò*, *perderò*; *parlerei* < *-arei*, *perderei*; Maiden 1995: 44, 159).

Plausible as a phonological account might be in the abstract, however, there are reasons to rule it out here. For one thing, no such sound change affecting pre-tonic *e* is attested independently in the relevant varieties, as it is in Florentine; indeed, the existence of both outcomes (*-ar-* and *-ir-*) scattered across the Occitan map would presuppose the existence of both raising and lowering effects which prevailed in different areas, but no such divergence is found outside the verb system. Moreover, there is in fact a third innovative development also attested for verbs of the *plànher* subtype: rather than taking on a different theme vowel in the future and conditional, they may lose it altogether, thus patterning with the verbs in *-re* which represent the majority of the third conjugation (e.g. 3sg *plandra* in Pézenas: Mâzuc 1899, cited in Esher 2012: 182). Again, the loss of an unstressed *e* in this position (followed by further changes in the resulting consonant clusters) is not intrinsically unlikely as a potential result of sound change. However, the presence of this innovation alongside the other two makes even more remote the possibility that these forms are all the automatic result of processes taking place below the level of the morphology.

The possibility that sound change played a role in these developments is further weakened by the more fine-grained morphological data brought to light by Esher from the unpublished findings of the dialect surveys *Atlas linguistique du Languedoc Oriental* (*ALLOr*, Boisgontier 1981-86) and *Atlas linguistique du Languedoc Occidental* (*ALLOc*, Ravier 1978-93). As she points out (p. 183), over the area where the relevant data have been collected for several different verbs, very few survey points are in fact found in which every verb in *-er* is treated alike with respect

to the identity of the theme vowel in the future/conditional forms. Much more commonly, there is variation at the level of the individual lexical item. Thus in one group of survey points in Gard, *córrer* takes on the vowel *i* in these forms while *mólzer* takes on the theme vowel *a*. Such differential treatment is hard to explain as the result of sound change. It is particularly telling that these local differences in behaviour are found even between stems which are very similar in phonological shape, both historically and in the present state of the language. Thus attestations of *náisser* (< **nāscere*) with future and conditional forms in *-ar-* are found scattered over the region covered by *ALLOr* (specifically in Aude, Hérault and Ardèche); by contrast, the formally similar verb *conéisser* (< *cognōscere*) shows forms in *-ar-* at one survey point only, and instead is much more likely to show loss of the theme vowel altogether in the future and conditional.

This divergent treatment even of such similar stems underlines the fact that these developments cannot have taken place on the basis of a sound change targeting any particular sequence found in the future and conditional of the relevant verbs; nor is there any way to interpret the anomalous paradigmatic behaviour of these verbs as phonologically motivated in synchrony. Instead, the widespread emergence of innovative forms in the future and conditional of Occitan verbs in unstressed *-er* must be recognized as a genuinely morphological phenomenon from the outset. In those instances where a given verb of this kind does not retain conservative *-erai/-eria* (or the equivalent) in the future and conditional in a particular variety, but instead appears to make use of a different theme vowel in these forms or to make no use of a theme vowel at all, this is because these parts of the paradigm have been remodelled directly on the basis of existing morphology which already characterized other classes of verbs in the lexicon.

An account for this process in terms of morphological factors (namely the characteristics of the conjugations and their lexical distribution) is readily available. As in all Romance languages, the continuants of the third and second conjugations of Latin make up the least well-populated and productive inflectional class in the Occitan verb lexicon; and the verbs in unstressed *-er*, with their distinctive future and conditional forms, further belong to a minority within this class. Accordingly, the distinctive future/conditional formation with theme vowel *e*, restricted to this group of verbs only, is substantially outnumbered by every other type. At the same time, the material difference between the original formation in *-er-* and its more lexically frequent counterparts is only slight, consisting only of the modification of an unstressed theme vowel. Under these circumstances, it might come as no surprise to find that speakers may remodel a less familiar formation, albeit a transparent one, on the basis of another which is far better represented over the language as a whole.⁴⁷ What is more, each of the morphological innovations attested by verbs of this type finds a motivation in different existing features of the verb system. Adoption of the theme vowel *a* in place of *e* aligns the future and conditional of these verbs with the most numerous and productive conjugation in the language; on the other hand, as noted above, the morphology of the third conjugation is more closely affiliated with that of the second (Esher 2012: 160, citing Dalbera 1994: 619f.). Loss of the theme vowel *e* altogether, meanwhile, alters the shape of the verb forms involved most substantially, requiring some phonological repair at the juncture of the stem-final consonant and the *r* of the future/conditional (cf. *plandrà* from *plànher* above), but morphologically it represents the least radical change, as it brings these verbs into

⁴⁷ It is notable that heteroclisis of this kind does not appear to take place in Gascon varieties, where the infinitive ending *-er* and associated future and conditional forms with theme vowel *e* are regular for the whole of the third conjugation rather than just a minor subset of the verbs belonging to it.

line with others of the third conjugation, whose morphology they already share in every other part of the paradigm except the infinitive.

Where the inherited third conjugation forms are replaced by those proper to the first or second conjugation, this is evidently incursion of innovative morphological material into the paradigm, as was seen above for the nouns treated by Karatsareas 2011a; and the paradigms which arise from this change can be described as heteroclite. Not every third conjugation verb is affected, and usually not even every third conjugation verb in unstressed *-er* behaves in the same way in a given dialect, if it shows innovative conjugational behaviour at all. This means that future and conditional formations in *-ar-* and *-ir-* retain their unambiguous affiliation with the first and second conjugations, and yet there exist individual verbs, otherwise showing third conjugation morphology, which also make use of these formations. Esher (p. 180) comments that what makes it possible to identify (synchronic) heteroclis in these cases is the fact that ‘the patterns... encountered apply to so few lexemes and with such little systematicity that they remain a minority’. This is essentially true, although even if innovative conjugational markers were used entirely systematically but by only a small number of the available verbs in the class involved, we would still readily wish to refer to this as heteroclis. Only at the point where most or all verbs in unstressed *-er* in a given variety reliably showed futures and conditionals in *-ar-* (for example) would it make more sense to regard *-ar-* as equally proper to both the first conjugation and the class in unstressed *-er*.⁴⁸

It is less clear whether the label of heteroclis is appropriate for the third type of innovation shown by these futures and conditionals, where the theme vowel *e* is omitted to give forms modelled on the behaviour of the third conjugation verbs with

⁴⁸ It is unclear whether any such varieties exist. Cf. what is practically the reverse case in Italian futures and conditionals such as *parlerò*, *parlerei*, as mentioned just above.

infinitive *-re*. The validity of the term depends on the extent to which the distinction between these two subclasses of the third conjugation can be considered a morphological one in the first place, a question which Esher notes but does not pursue; though the fact that remodellings such as *planherai* to *plandrai* are possible at all entails that there is no live phonological principle making it impossible also to produce an infinitive without a theme vowel, which would collapse the two subsets of the third conjugation in the given varieties: this suggests that a morphological principle is still at work keeping the two separate.

However this may be, Esher's study clearly demonstrates the prevalence of one general type of heteroclisis in Occitan, which takes different specific forms from variety to variety and from word to word. It also provides evidence of a trigger for heteroclisis which implicates phonology 'indirectly', in a somewhat similar way to that seen in the case treated by Karatsareas above. There, the adoption of alien inflectional morphology in certain parts of the paradigm seems to have taken place in an environment which could be characterized in terms of the shape of the relevant inherited forms, but not as a result of sound change; instead, the problematic nature of the inherited accentual pattern of a particular set of nouns apparently led to a tendency towards the observed morphological shift. In the Occitan case, meanwhile, the future and conditional forms of the IIIb verbs have not undergone sound change or become 'unpronounceable' on phonological grounds; the inherited forms involved are simply *unfamiliar* in comparison with the more common types of future and conditional existing in the relevant Occitan varieties, and the situation is such that only a set of verbs characterized by particular stem shapes are affected.

The remainder of this thesis looks in more detail at several further instances of heteroclisis in the Latin/Romance family, after a historical introduction to its

fundamental conjugational makeup, which has been passed over swiftly up to this point. As will be seen from the case studies which follow, sound change alone is apparently sufficient to give rise to heteroclisis; on the other hand, it may also arise as the result of more complex chains of morphological development than have been seen so far.

3. Heteroclisis in the Latin verb

3.1 Introduction

This chapter considers heteroclisis in the Latin verb system in its diachronic and synchronic aspects. We have seen above that the concept of heteroclisis is relevant to the Latin noun system, as it can be invoked to characterize the irregular inflectional behaviour of *domus*; here I investigate the inflectional behaviour of several groups of Latin verbs which can be observed to show the same phenomenon, but which have not previously been viewed in these terms.⁴⁹

The chapter takes the following form. I begin with an overview of the structure of a typical Latin verb paradigm, together with an account of the four basic conjugational classes and the role they play in determining the morphological material with which the cells of the paradigm are filled for different verbs. This is followed by an introduction to the nature of the proto-Indo-European (PIE) verb system, which comparative reconstruction across the Indo-European languages shows to have differed markedly from the system seen in Latin (and the Sabellic languages, which make up Latin's sister branch within Italic). After sketching the developments which led from the PIE verb system to its Latin descendant, as far as is

⁴⁹ Heteroclisis is, in fact, proposed as relevant to the Latin verb system in Walther 2011, Sagot and Walther 2011; however, these treatments make use of the concept of heteroclisis specifically to analyse a phenomenon generally regarded as the classic instance of *deponency*. Across most of the paradigm, the Latin 'deponents' show passive morphology although they retain active syntax, e.g. (*eum*) *sequor* 'I follow (him)', cf. passive *dūcor* 'I am led'; however, they also possess present participles with normal active morphology (e.g. *sequēns* 'following', cf. *dūcēns* 'leading'). While voice is normally taken as an inflectional category in Latin, the above analyses treat it as derivational, on the grounds that its semantic effects may be somewhat unpredictable (Kiparsky 2005), and thus treat what has traditionally been seen as dedicated active and passive morphology as belonging to two separate, purely formal *conjugations* orthogonal to those discussed here. Accordingly, for them, Latin 'deponent' verbs in fact bring together morphology from two conjugations in a single paradigm, and are hence not deponent but heteroclitite. With Spencer (2013: 97f.), I find this approach to Latin voice untenable, and do not consider the notion of heteroclisis appropriate to this case.

necessary to explain the background to the studies offered here, I turn to particular instances of heteroclisis.

The first case treated in this chapter concerns verbs such as *capiō* ‘take’, whose paradigm appears to be distinctively affiliated with both the third and fourth conjugations in different cells. As will be discussed below, the opposition between these two conjugations is primarily valid only with reference to those parts of the paradigm which make use of the verb’s present (or *inflectum*) stem.⁵⁰ The second instance of heteroclisis identified here, meanwhile, does take the whole paradigm as its scope: verbs of this kind, such as *domō* ‘tame’, have paradigms which are split neatly between distinctively first conjugation and non-first conjugation behaviour. In these cases, all cells instantiating the verb’s present stem pattern together, usually in opposition to the entire rest of the paradigm. In a third section I treat a group of verbs (e.g. *petō* ‘seek’) showing the same paradigmatic split between the *inflectum* and elsewhere, but this time implicating the third and fourth conjugations: it is notable that this *petō* type and the *capiō* type are able to coexist and remain distinct in Latin. In all these cases, discussion of the synchronic status of the inflectional behaviour at issue is accompanied by an investigation of its diachronic origins. A final section summarizes and concludes.

3.1.1 *The Latin verb paradigm and the (present-stem) conjugations*

As is generally the case across the IE family, the verb paradigm of Latin is by far the largest and most complex morphological paradigm found in the language. Finite verb forms are sensitive to distinctions of person, number, mood, tense and voice, and each verb also possesses a number of non-finite forms, including infinitives and participles which are distinguished for tense and voice (as well as the expected

⁵⁰ See section 3.1.1 immediately below for this term.

adjectival categories of gender/number/case for the participles). These inflectional distinctions are largely realized suffixally, but may also involve weak or strong stem suppletion, and a great deal of cumulative exponence is found. Crucially for our purposes, verbs fall into separate inflectional classes, conventionally known as conjugations, according to their inflectional behaviour particularly in certain parts of the paradigm. The logic behind the division of verbs into these classes will be discussed shortly.

As an introduction to the Latin verb system, Table 3.1 gives a partial paradigm of the verb *moneō* ‘warn’⁵¹ (based on the layout of the paradigm given in Meiser 1998: 178), which illustrates the principal finite and non-finite inflectional values available to a typical Latin verb and the way in which these are realized by the morphology (though for caveats involving slight variations in behaviour across the conjugations, see immediately below). Only the active portion of the finite paradigm is shown here. All finite formations are exemplified by the 1sg except for the imperatives, for which the 2sg is given.

Table 3.1 Outline of the paradigm of *moneō* ‘warn’

	Indicative	Subjunctive	Imperative	Infinitive
infectum				
Present	<i>mone-ō</i>	<i>mone-am</i>	<i>monē</i>	<i>monē-re</i>
Imperfect	<i>monē-bam</i>	<i>monē-rem</i>		
Future	<i>monē-bō</i>		<i>monē-tō</i>	
perfectum				
Perfect	<i>monu-ī</i>	<i>monu-erim</i>		<i>monu-isse</i>
Pluperfect	<i>monu-eram</i>	<i>monu-issem</i>		
Future perfect	<i>monu-erō</i>			
Further non-finite forms				
Based on the present stem:				
pres. active participle <i>monē-ns</i> , gerund <i>mone-ndum</i> , gerundive <i>mone-ndus</i>				
Based on the third stem:				
perf. passive participle <i>monit-us</i> , fut. active participle <i>monit-ūrus</i> , supine <i>monit-um</i>				

⁵¹ In line with tradition, Latin verbs are cited here in the 1sg present indicative active form; sometimes the infinitive ending is also provided where this disambiguates the conjugation of the verb.

Certain features of this paradigm require comment. The tendency towards cumulative exponence is obvious here, and may be clearest in the expression of the distinction between indicative and subjunctive where both exist: evidently there is no single morph realizing either of these moods. As this is a consistent characteristic of Latin verb morphology, I have attempted only the bare minimum of morphological segmentation necessary to distinguish broadly between stems and inflectional endings, though as will be seen, even this division can be problematic. In terms of the shape of the paradigm itself, perhaps most notable is the fact that the inflectional dimension which could be labelled as ‘tense’ in this table has an internal structure of its own: thus the six Latin tenses given here fall into two parallel groups, the opposition between imperfect, present and future within the part of the paradigm known as the *infectum* being mirrored by an equivalent opposition between pluperfect, perfect and future perfect, which together make up the *perfectum*.⁵²

The functional distinction between these two parts of the paradigm is not easy to characterize, as each of aspect and relative tense⁵³ has been proposed to the exclusion of the other; it is likely, in fact, that both may be involved depending on the linguistic context. However, the distinction clearly plays an important role in the formal articulation of the paradigm. All active finite forms in the *infectum*, together with the associated present infinitives and some other non-finite forms (present active participle, gerund and gerundive), are built on the same ‘present stem’, in this case *monē-*: this shows only the slightest of contextual modifications, its final vowel undergoing regular shortening before a sequence of nasal + dental stop (as in *monendus/monendum*) or before a following vowel (*moneō* itself, and *moneam*), both

⁵² These names are already used to characterize the two formal parts of the finite paradigm in Varro, *De Lingua Latina* (first century BC).

⁵³ See e.g. Pinkster (1990: 229-32, favouring anteriority) and Touratier (1994: 114-16, favouring perfectivity) on the meaning of the *perfectum*.

cases resulting from fairly recent sound changes which had effects outside the verb system too.⁵⁴ The active *perfectum* forms, meanwhile, make use of the ‘perfect stem’ *monu-*. The distinction between these two halves of the finite paradigm is also salient in the passive, where synthetic *infectum* forms again make use of *monē-* whereas all *perfectum* forms are supplied by periphrases involving the participle *monitus*, which is built on the ‘third stem’ *monit-*.⁵⁵ This third stem is never found in synthetic finite forms: these are based solely on the dedicated present and perfect stems. On the other hand, the third stem also provides the future active participle and the supine (a secondary infinitive or verbal noun most commonly used after verbs of motion), and may serve as the basis for derived forms outside the paradigm (e.g. the agent noun *monit-or*). Within the verb paradigm, zones making use of these three stems can thus generally be taken to constitute coherent domains for the application of particular PSCs: the traditional ‘principal parts’ of Latin verbs represent a potential set of *reference forms* (cf. section 2.2.2) which play a role in the operation of these PSCs.

All verbs in Latin, with a very few exceptions which are also recognizable as morphologically anomalous on other grounds, show the same predictable distribution of three distinct stems over the finite paradigm as is seen here in the case of *moneō* (Meiser 1998: 179). However, it does not follow that there is an equally predictable relationship between the *forms* which instantiate those stems. In the current instance, *monē-*, *monu-* and *monit-* are straightforwardly related in shape, and they can be treated synchronically as produced by the suffixation of opposing stem formants on the basis of a single lexical root. But this is not the case for all verbs. The verb

⁵⁴ Sihler (1995: 77f., 80). The latter change survived in the form of a ‘live’ phonological constraint on the language: the synchronic generalization that long vowels do not appear in prevocalic position was made by Roman grammarians themselves, cf. Quintilian *Institutio Oratoria* 12.10.57, *vocalis ante vocalem corripitur* ‘a vowel shortens before a vowel’.

⁵⁵ The name ‘third stem’ is provided by Aronoff 1994: he identifies this stem as an example of a ‘morphomic’ – purely morphological – entity, as the forms built on it have nothing in common besides this agreement in shape.

mordeō ‘bite’, with present stem *mordē-*, has perfect stem *momord-* and third stem *mors-*; the verb *videō* ‘see’ has the stems *vidē-*, *vīd-*, *vīs-*. In each case the formal relationship between the three stems is different, illustrating the fact that there is no general expectation that the identity of any single Latin verb stem will allow for the prediction of any other. Though this claim will be qualified shortly, the point stands that speakers of Latin could be required to store three (synchronously) mutually unpredictable stems for a single verb, which between them account for the forms found in all the cells of that verb’s paradigm.

The shape taken by one of these stems in particular, the present stem, has ramifications for other aspects of the inflection of a given lexeme. While it is possible, on historical grounds at least, to distinguish a number of formal groups within the variety of perfect and third stems available in Latin (e.g. *momord-* represents the ‘reduplicated perfect’, while *vīd-* represents the ‘lengthened-vowel perfect’), these distinctions are morphologically inert: the inflectional endings required in the relevant parts of the paradigm are not dependent on the shape of the stem. In the *inflectum*, however, a verb’s inflectional behaviour is determined by the identity of its stem-final (or ‘theme’) vowel. This is illustrated below by means of partial paradigms from the *inflectum* of the verbs *cūrō* ‘care for’ and *regō* ‘rule’, whose infinitives *cūrā-re* and *rege-re* can be taken to display this theme vowel: both of these contrast in some respects with what was seen for *monē-re* above.

Table 3.2 Some *inflectum* forms of first conjugation *cūrō -āre* ‘care for’ and third conjugation *regō -ere* ‘rule’

	present indicative	present subjunctive	future indicative	present indicative	present subjunctive	future indicative
1sg	<i>cūrō</i>	<i>cūrem</i>	<i>cūrābō</i>	<i>regō</i>	<i>regam</i>	<i>regam</i>
2sg	<i>cūrās</i>	<i>cūrēs</i>	<i>cūrābis</i>	<i>regis</i>	<i>regās</i>	<i>regēs</i>
3sg	<i>cūrat</i>	<i>cūret</i>	<i>cūrābit</i>	<i>regit</i>	<i>regat</i>	<i>reget</i>
1pl	<i>cūrāmus</i>	<i>cūrēmus</i>	<i>cūrābimus</i>	<i>regimus</i>	<i>regāmus</i>	<i>regēmus</i>
2pl	<i>cūrātis</i>	<i>cūrētis</i>	<i>cūrābitis</i>	<i>regitis</i>	<i>regātis</i>	<i>regētis</i>
3pl	<i>cūrant</i>	<i>cūrent</i>	<i>cūrābunt</i>	<i>regunt</i>	<i>regant</i>	<i>regent</i>

The paradigms differ in several ways: notably, *cūrō* patterns with *moneō* (and against *regō*) with regard to its future formation in *-bō*, while *regō* patterns with *moneō* (and against *cūrō*) with regard to its subjunctive formation in *-am*. Meanwhile, in both *cūrō* and *regō* the theme vowel is not reflected in the 1sg present, unlike in *moneō*. In fact, in the paradigm of *regō* a theme vowel *e* can hardly be said to surface at all as such, unlike *ā* in the paradigm of *cūrō*. All this makes the point that the identity of the theme vowel determines not only what endings are selected, but how they interact with the theme vowel itself, to the extent that in some cases it makes no sense to treat the two as separable units.

The three inflectional classes illustrated by *cūrō -āre*, *moneō -ēre* and *regō -ere* constitute the first, second and third conjugations of Latin respectively; the behaviour of the fourth conjugation (e.g. *audiō* ‘hear’), whose theme vowel is *ī*, will be treated alongside that of the third in section 3.2 below. These four conjugations are all that are usually recognized for Latin, and, leaving aside those verbs which will be the subject of 3.2, only a small number of very high-frequency verbs (such as *edō* ‘eat’ and *volō* ‘want’, also irregular in other ways) have present stems which cannot be accommodated in this categorization.

The influence of conjugational identity, which fundamentally concerns the *inflectum*, is not entirely restricted to this portion of the paradigm, however. Notably, with very few exceptions, verbs belonging to the first conjugation – with theme vowel *ā* – preserve this theme vowel in their *perfectum* and third stems, which take the forms *-āv-*, *-āt-* respectively (e.g. perfect *cūrāvī*, supine *cūrātum*). To an extent the fourth conjugation behaves analogously, giving *-īv-*, *-īt-* in these stems (e.g. *audīvī*, *audītum*), although this behaviour is significantly less reliable and not characteristic of many of the most commonly found fourth conjugation verbs. The

second and third conjugations, meanwhile, are characterized by a multitude of different *perfectum* and third stem forms (Leumann 1977: 603f.)

The foregoing overview of the Latin verb system provides the synchronic context in which the instances of heteroclisis discussed in sections 3.2 to 3.4 below are to be regarded. I now turn to the diachronic background to this system. After the sketch of the PIE verb which follows, I briefly illustrate how the state of affairs reconstructed for PIE developed to give the regularities of Latin conjugational behaviour presented in their fundamentals just above, reserving more detailed treatments of individual questions to the relevant sections. As will become clear, however, not only these regularities but also much of the heteroclitite behaviour of interest to us here can be traced back to features of the PIE system. In particular, some of the peculiar features of heteroclisis of the ‘*domō* type’ treated in section 3.3 make more sense in the light of our understanding of the PIE verb. On the other hand, the behaviour of Latin verbs of the ‘*capiō* type’ has been taken to point to the existence of a particular formation in PIE, a claim which I investigate in section 3.2.

3.1.2 An overview of the PIE verb

I begin with a description of the PIE verb system. For this overview I rely principally on the ‘Cowgill-Rix’ reconstruction,⁵⁶ as is presented for example by Sihler (1995: 442-515), Ringe (2006: 24-41) and especially Rix 2001 (*Lexikon der indogermanischen Verben*, 2nd edition, henceforth *LIV*²), and commonly taken to represent a proto-stage of the Italic system (e.g. Bock 2008: 17-19). This is based on the analysis of a vast array of evidence from all the attested branches of IE, chiefly

⁵⁶ This reconstruction in fact targets only a large clade within Indo-European, called West IE by Ringe (2006: 5); this does not include Anatolian and Tocharian, which split off earliest from the true proto-language. Nonetheless, as Italic shares this clade with practically all the IE languages, and the best evidence for the reconstruction presented is provided by Greek and Indo-Iranian, both with verb systems utterly unlike that of Latin, it clear that to begin at this point is to miss nothing with regard to the specifically Italic developments which we will be interested in.

Greek, Indo-Iranian and Italic itself, but for reasons of space I will be able to provide little besides my own selective version of the results of this analysis.⁵⁷ the presentation of the evidence itself would take us too far afield. Inevitably, I also leave out a great deal which can be more or less securely reconstructed but is not germane to the topics treated in this thesis; this presentation is largely limited to treating those features of the PIE verb most relevant to the origins of the Latin conjugations and the three-stem structure of the Latin paradigm.

The verb system of (late) PIE was fundamentally underpinned by distinctions of *aspect*. The paradigm of a basic or ‘primary’ (i.e. underived)⁵⁸ verb was articulated according to an opposition between stems expressing durative/imperfective, punctual/perfective and stative aspect, which have acquired the labels *present*, *aorist* and *perfect stem* respectively after the traditional names of their descendants as found in Ancient Greek, where the distinction largely survives both formally and functionally. In morphological terms the more fundamental split here is between the perfect stem and the others, i.e. between stative and eventive verb forms: one sign of this is the fact that the present and aorist selected from a common pool of inflectional markers distinct from those used by the perfect. All three stems had an imperative, but it is likely that only eventive stems inflected for the further moods subjunctive and optative, which were marked by specific suffixes preceding the personal endings (which marked person, number and voice). Both of these expressed *irrealis* modality, but of different kinds: the optative was used to

⁵⁷ My presentation of the PIE state of affairs follows Sihler and Ringe, while for stem types and the behaviour of individual verbs I generally follow the reconstructions given by *LIV*². For the evidence and argumentation which underlies current views of the behaviour of the PIE verb, the works listed above may be consulted along with many others, including Fortson (2004: 81-101); Clackson (2007: 114-56); Meier-Brügger et al. (2010: 295-320); Szemerényi (1996: 230-338), and I direct readers to these comprehensive overviews rather than providing separate references for the various components of the picture outlined in this section. Presentations of the PIE verb which diverge more substantially from that given here include Watkins 1969; Jasanoff 2003; Beekes (2011: 251-88).

⁵⁸ See below (p. 144) for the relevance of the distinction between basic and derived verbs.

express potentiality or the wishes of the speaker, and the subjunctive to characterize an event as somehow uncertain or less than totally settled from the speaker's point of view. The subjunctive could thus be used to refer to future events.

There is no general distinction between present and future tense in the PIE verb, however, and the category of tense as such is rather weakly developed in PIE. Only the indicative shows a distinction between present and past tenses, and this only in the present stem system: formally this distinction was encoded in the opposition between 'primary' and 'secondary' personal endings, which gave present and imperfect verb forms respectively. The aorist indicative, meanwhile, was found with secondary endings only, meaning that the theoretically available but logically problematic combination 'punctuality with present reference' was not exploited in any extended meaning which might be conceivable for it.

Present, aorist and perfect between them exhaust the aspects available to a basic verb at this stage in the development of PIE. Any finite form will belong to one of these three series; there also exist non-finite forms, the participles, associated with each of them. The use of one aspect or another determines both the further inflectional distinctions which the verb form may realize and the shape of the stem itself, before the further affixation of participial or modal marking or of the personal endings directly. But the behaviour of individual verbs in forming their aspect stems contains a great deal that is unpredictable. First of all, it is a striking feature of the PIE system that there is no guarantee that a given verb will in fact possess all three aspect stems available to it in principle. In some cases there is certainly evidence for all three dating back to the proto-language: Ringe (2006: 34) offers the example of the root **g^wem-*, which formed the aorist **g^wém-* 'step', the present **g^wm-ské-* 'be

walking’,⁵⁹ and the perfect **g^we-g^wóm-* ‘have the feet in place’. But this is far from the norm. Some verbs possessed only one aspect stem, e.g. present **wés-* ‘be wearing’, or only two, e.g. present **dék-* ‘be pulling’ / aorist **dék-s-* ‘pull’. The aspect stems, in effect, behaved to an extent more like the results of derivation based on the root than inflectional realizations of it. It seems clear that this situation reflects an earlier form of the system, based on lexical *Aktionsart* rather than aspect proper: I return to this below.

Secondly, PIE verbs were not uniform in the morphological processes which they employed in the production of these aspect stems. All stems of a given basic verb were formed on the basis of a monosyllabic verbal root, whose syllabic nucleus seems always to have been the vowel *e*. This root encoded the fundamental semantics of the verb, and its existence as a genuine morphological element manipulable by speakers, at least at the period in question here, is demonstrated by the fact that verbal nouns and adjectives were derived directly from it and not via the mediation of any of the aspect stems. But of these stems, the perfect is the only one which consistently stands in the same formal relationship with the root:⁶⁰ as seen in the case of **g^we-g^wóm-* above, the perfect stem is realized by reduplication of the initial consonant (or cluster) with *e* as the supporting vowel. The vocalism of the root syllable is changed to *o*, in the indicative singular at least; but in fact this vowel alternates predictably between *o*, *e* and zero over the perfect paradigm, a purely morphologically motivated phenomenon known as *ablaut* and found extensively

⁵⁹ In this section only, the English progressive is used where necessary to signal the durative nature of the PIE present stem.

⁶⁰ The lone exception to this generalization is *wóyd-* ‘know’, which shows no trace of reduplication.

throughout PIE.⁶¹ Reduplication and ablaut, as well as affixation, are standard processes of stem formation elsewhere in the verb system too.

The aorist and (especially) the present stems, meanwhile, are formed more variously. On the basis of comparison across the daughter languages, *LIV*² attributes to PIE three aorist formations and as many as 21 distinct present formations, and only those of greatest importance to the development of the system into Italic can be presented here. What is more, these eventive stems fall into two morphological types, which cut across the functional distinction between present and aorist. As was stated above, present and aorist stems pattern together in many ways in terms of their inflectional exponence; but this statement comes with the qualification that stems of both kinds are divided into the two form classes known as *thematic* and *athematic*, which show strikingly different morphological behaviour in some important respects. The basic formal distinction between the two is that thematic stems possess a final vowel segment, which shows regular ablaut between *e* and *o* depending on its morphological context: this is the single *thematic vowel* of PIE, which is also found as a structural element in the inflectional system of nouns and adjectives and is continued in, among other things, the inflected forms of the Latin second declension (nominative *-us* < *-o-s*, vocative *-e*). Athematic stems lack this element, and always end in a consonant.

The presence of the thematic vowel has two principal effects. Firstly, as elsewhere in PIE morphology, thematic stems are always invariant in their shape, except for the ablaut undergone by the thematic vowel **e/o* itself: word accent remains fixed, whether on the thematic vowel or not, and the stem shows no ablaut

⁶¹ It seems clear that, at least at a certain stage in the development of the language, the vowel *o* did not exist in PIE except as a conditioned variant of *e*, and thus the two are often treated as a single item for descriptive purposes. The conditions which determined the appearance of *o*, or of no vowel at all, in place of the default *e* were most likely phonological to begin with, but the comparative method takes us back only to a stage where the distribution of these vowels has already been morphologized.

otherwise. This is not usually the case in athematic stems, where accent shift and ablaut effects have the potential to be much more wide-ranging: in one common pattern, both the position of the accent and the ‘ablaut grade’ of the stem differ between singular and plural in the indicative (e.g. the aorist stem $*g^wém-$ ‘step’ given above actually appears as ‘zero-grade’ $*g^wm-$ in the indicative plural forms, which transfer the accent to the personal endings). The second effect is that the formal class of the stem can, in a few cases, determine the choice of inflectional marker which follows, making the opposition between thematic and athematic in PIE analogous to the opposition between the Latin conjugations as seen in the previous section. The most striking difference between the two is seen in the 1sg primary ending. In athematic verbs, this is $*-mi$, which reflects the addition of a so-called *hic-et-nunc* marker $-i$ to the secondary ending $*-m$ (compare the other singular endings: primary $-si -ti$, secondary $-s -t$). But while thematic verbs share this secondary ending $*-m$ (which combines with the thematic vowel to give $*-o-m$), the corresponding primary form is not $**o-mi$ but $*-o-h_2$.⁶²

A small number of verbal roots formed thematic aorist stems in PIE. The only type of this kind recognized in *LIV*² is the reduplicated thematic aorist, exemplified by $*wé-wk^w-e/o-$ from the root $*wek^w-$ ‘say’. Two athematic formations were overwhelmingly more common. The root aorist, as the name suggests, makes use of no formal material at all to characterize the aorist stem: the stem is simply the root itself, with the appropriate ablaut grade and accent. The aorist $*g^wém-$ (/ $*g^wm-$) ‘step’ seen above is of this type. The sigmatic aorist, meanwhile, adds the athematic

⁶² Three segments conventionally labelled $*h_1$ $*h_2$ $*h_3$, the so-called ‘laryngeals’, can be reconstructed, which act phonotactically as obstruents in PIE but are generally continued as vowels in the daughter languages (if they survive at all). Often they are found to lengthen or colour a vowel they are in contact with: for example, the sequence $-oh_2$ mentioned here is the source of Latin 1sg $-ō$. If not in contact with a vowel, all three give Latin *a* (Sihler 1995: 99f.).

suffix *s* to a ‘lengthened-grade’ root and is exemplified by **dēwk-s-* (/ **dēwk-s*) from the root **dēwk-* ‘pull’.

The present stems of PIE are much more variable than this. Some verbs form entirely uncharacterized presents, which, like the root aorists above, make use of no extra morphological material at all beyond the root. The stem **dēwk-* (/ **duk-*) ‘be pulling’ already seen above is a root present of this kind. These exist alongside uncharacterized counterparts belonging to the thematic form class, such as **b^hér-e/o-* ‘be carrying’, and in fact the latter seems to have been a relative newcomer (it is practically unattested in Anatolian) and a productive type in PIE, gradually attracting roots which had originally produced root presents. Reduplicated presents are also found in both athematic and thematic types, and distinct formations existed which made use of *i* and *e* in the reduplicating syllable: examples are athematic **ké-koyd-* (/ **ké-kid-*) ‘be calling’ from **keyd-*, thematic **sí-sd-e/o-* ‘be (in the process of) sitting down’ from **sed-*. Here too it is clear that the thematic forms were gaining ground at the expense of the athematic forms already in PIE.

One important present stem formation is morphologically unique within PIE. It involves the infixation of **-né/n-* before the final consonant (or cluster) of the root, giving, for example, present stem **li-né-k^w-* (/ **li-n-k^w-*) ‘is leaving behind’ from the root **lik^w-*. This is the only infixed present stem formant, and indeed the only infix reconstructed anywhere in the morphology of the proto-language. Suffixes, both athematic and thematic, were more numerous. A few verbs made use of an athematic suffix *-u-*, such as **wél-u-* (/ **wl-ú-*) ‘be rolling’ from **wel-*. The many thematic present-stem suffixes included *-d^he/o-* (**kél-d^he/o-* ‘be driving’ from **kel-*), *-ské/ó-* (as in **g^wm-ské/ó-* ‘be walking’), and *-ye/o-* (**spék-ye/o-* ‘be looking’ from **spek-*).

Like the uncharacterized thematic presents, the type in unaccented *-ye/o-* seems to have become productive relatively late.

All these basic stem types are continued into Italic, although it is clear that the structural distinctions between them have been obscured a great deal by intervening developments to give rise to the very different conjugational system seen above for Latin. But their characteristics as presented above are also indicative of the nature of the PIE verb system at an earlier stage. It is notable and highly revealing that we can reconstruct both root aorists and root presents, i.e. athematic aorist stems and athematic present stems which were formed by the addition of no material whatsoever to the root, although no single verb root could form both. As all athematic eventive stems drew on the same inflectional affixes and personal endings (save that aorists could not take the ‘primary’ endings with present reference), this means that there was no difference in morphological *structure* between an aorist secondary indicative such as 3sg **g^wem-t* ‘stepped’ and a present secondary (i.e. imperfect) indicative such as 3sg **dewk-t* ‘was pulling’: the only difference lay in their morphological *relationship* with other forms of the same root, which required overt morphological material to signal present and aorist aspect respectively. This is our evidence for supposing that aspect was originally not grammatical but a semantic feature inherent to individual lexical roots (*Aktionsart*): the various aspect-stem formants would then reflect earlier derivational affixes which, with weakening of their earlier semantics, were reinterpreted as marking aspectual distinctions only.

The idea that an earlier *Aktionsart* system lies behind the aspect system reconstructed above helps to explain some of its idiosyncrasies. The failure of many verbs to produce all available aspect stems makes more sense if the system we reconstruct had recently crystallized out of a stage where there was originally no

grammaticalized aspect opposition, but only individual stems each (more or less incidentally) bearing its own aspect value. The large number of functionally equivalent present-stem formations seen above is also accounted for on the assumption that this equivalence arose secondarily as the aspect system established itself.

In support of this is the fact that we can reconstruct several verbal formations which clearly still make a semantic contribution, and which are taken as markers of deverbative *derivation* in descriptions of PIE, but which have been accommodated within the aspect system in that the forms they create belong definitively to one aspect or another. These give an impression of the earlier status of the purely aspectual stem formants discussed so far. Deverbative formations aligned with the present aspect include the desiderative in **-se/o-*, e.g. **wéyd-se/o-* ‘want to see’ from **weyd-* ‘catch sight of’, and a formation which could have either causative or iterative semantics: this used a suffix **-éye/o-* and required *o* in the root syllable, e.g. (causative) **mon-éye/o-* ‘be warning’ from **men-* ‘think’, and (iterative) **b^hor-éye/o-* ‘be carrying around’ from **b^her-* ‘carry’. Finally, an athematic deverbative in **-eh₁-* was also found (labelled the ‘essive’), which appears to have a meaning related to stativity, although it is unclear whether it originally operated as a present or an aorist stem, e.g. **h₁rud^h-éh₁-* ‘be red / become red?’ from **h₁rewd^h-* ‘red’.

Meanwhile, present stems also existed which showed the results of deadjectival and denominal derivation. Stems of the former type were athematic, and involved the suffixation of **-h₂-* to the adjectival stem (e.g. **néw-e-h₂-* ‘be renewing’, from **néw-e/o-* ‘new’). The latter made use of the thematic suffix **-ye/o-* added to the stem of the original noun. Interestingly, however, a pattern developed whereby thematic nouns, which might have been expected to produce denominal

verbs in **-e-ye/o-*, in fact made use of the slightly different sequence **-eh₁-ye/o-*. In Latin terms this leads to the situation whereby the first conjugation produces denominal verbs not only to first declension nouns (in $\bar{a} < *eh_1$), e.g. *cūrāre* above from *cūra* ‘care’, but also second declension nouns, which reflect the thematic vowel, e.g. *dōnāre* ‘give’ from *dōnum* ‘gift’.

All these types of formation, adding semantic nuances rather than simply marking a stem as present, aorist or perfect, created *derived* stems, as opposed to the basic stems treated earlier. The distinction is important. Derived stems, usually integrated into the system as present stems, had no direct counterparts in the other aspects: thus there was no aorist counterpart to **néw-e-h₂-* ‘be renewing’ or **wéyd-se/o-* ‘want to see’. To the extent that deverbative stems ‘had’ aorists or perfects at all, these were simply the aorists and perfects built on the original root. Nor did verbal adjectives or verbal nouns exist in relation to these derived stems specifically, as they could only be formed on the basis of roots. The verbal adjective from the root **weyd-*, for example, took the form **wid-tó-*, and the verbal noun the form **wéyd-tu-*; these were the closest the system came to providing a verbal adjective and noun for derived **wéyd-se/o-* as well.

3.1.3 From PIE to the Latin conjugations

The PIE verb system just described evidently differs in many important ways from the Latin verb system exemplified by the behaviour of *moneō* above. This is no surprise, given the millennia which separate the proto-language from even our earliest attestations of Latin in the 7th/6th century BC. Not all of the differences are relevant to an investigation of heteroclisis in the Latin verb. I will point out only a few essential structural changes important to the composition of the Latin paradigm,

before turning to address the way in which the Latin conjugations emerge from the numerous stem types touched on above.

One crucial development concerns the functional distinction between the perfect and aorist stems of PIE, which has been lost in Latin; what is called the ‘perfect stem’ in Latin emerges as the result of this collapse. Formations originally proper to the two different aspect stems can both be observed in the Latin *perfectum*: thus Latin reduplicated perfects such as *pepigī* ‘I fixed’ < **pe-pag-ai* to *pangō* reflect the usual PIE perfect formation, while sigmatic perfects such as *dūxī* ‘I led’ < **dewk-s-ai* continue the PIE sigmatic aorist. The collapse of these two categories into one is a fairly late development in the history of Latin, and took place separately in Latin and the Sabellic languages, as can be seen from the fact that the two branches of Italic sometimes preserve opposing inherited formations as their representative of the perfect stem for a particular verb, e.g. Latin *dīx-* < aorist **deyk-s-* vs. Umbrian *dersic-* < perfect **de-dik-* for the perfect stem of ‘say’ (Meiser 2003: 73).

Meanwhile, the ‘third stem’ of the Latin verb comes about as the result of the incorporation into the paradigm of the verbal adjective and noun in *-tō-*, *-tu-* respectively (Sihler 1995: 620f., 613); the former appears as the perfect passive participle, and the latter as the supine. While these originally required different ablaut grades of the root, no such distinction survives in Latin (Weiss 2009: 445), and further formations were built on to the unified stem which thus resulted: e.g. perfect passive participle *ductus* < **duk-tō-*, supine *ductum* as if < **duk-tu-*, future active participle (of unknown origin) *ductūrus*, etc. Phonotactic constraints operating already within PIE led to assibilation of the cluster produced by these endings following a root-final dental: thus the Latin third stem often contains an *s* as opposed to a *t* (*vīsus* ‘seen’ < **wid-tō-*, *sēnsus* ‘felt’ < **sent-tō-*). This *s* could also be spread

to the third stem by analogy on the basis of a sigmatic perfect, signalling a developing morphological relationship between these two parts of the Latin paradigm with very different origins, e.g. *fixus* ‘fixed’ after perfect *fixī* (present *fixō*), cf. Leumann (1977: 615f). This morphological relationship was promoted by the special status of the present stem in the PIE verb system, whose ramifications for the morphology of the rest of the paradigm in Latin play an important role in the phenomena treated in this chapter and figure heavily in the discussions below.

The present stem formations of PIE are indeed continued in the Latin *inflectum*; however, it will be clear that the four-conjugation system of Latin does not correspond directly to any partitioning of the verb lexicon seen in PIE. I will start from one obvious point of continuity which does exist between the two systems. The thematic vowel, in PIE present stems which possessed it, showed an ablaut alternation between *e* and *o* over the paradigm: in the present indicative, *o* was found in the 1sg, 1pl and 3pl, and *e* in the remainder. This alternation is continued with perfect regularity (given sound changes widely attested otherwise for archaic Latin) in the behaviour of the theme vowel of the Latin *third* conjugation, where *o* survives in the 1sg *-ō*, has been raised by the following nasal in 3pl *-unt*, and has undergone ‘weakening’ to *i* in 1pl *-imus* (as is usual for short vowels in non-initial syllables),⁶³ while *e* has undergone the same weakening to *i* in *-is*, *-it*, *-itis*. The same *e/o* element can be seen in the present imperative endings (2sg *-e*, 2pl *-ite* with weakening), the present participial stem *-ent-*, and the gerund/gerundive stem *-end-* (also existing in the more archaic form *-und-*). These are important indications that the thematic present stems of PIE, in the absence of any phonological or morphological disruption, retain their common identity in the guise of the Latin present stems of the

⁶³ Weakening is discussed in greater detail in section 3.2.2 below.

third conjugation: this includes simple thematics such as **h₂eg-e/o-* ‘drive’ (> Latin *agō*) alongside characterized thematics such as **prk-ské/ó-* ‘demand’ (> Latin *poscō*) and **wéyd-se/o-* ‘want to see’ (> Latin *vīsō* ‘visit’), cf. Sihler (1995: 505f., 509).

At the same time, however, it is clear that many verbs which inflected athematically in PIE also make their home in the third conjugation. In part this is merely a continuation of the long-standing trend replacing root and reduplicated athematic presents with the corresponding thematic stems, which was well underway in the parent language. However, we can see that already in proto-Italic this process of thematization had expanded in scope to cover the nasal-infix presents as well: Latin *fingō* ‘fashion’ agrees with Umbrian **fiktu**⁶⁴ (i.e. /fingtu/ < **finge-tōd*) ‘he is to fashion’, both immediately descending from a remodelled **d^hing-e/o-* (perhaps already with initial fricative at the relevant stage) rather than the original athematic **d^hi-né-g- / d^hi-n-g-* (Meiser 1986: 82-84). A rather different kind of thematization process is seen in verbs such as Latin *vomō* ‘vomit’. Here, regular sound change affecting an original root present **wemh₁-* > **wema-* resulted in what looked like a thematic stem with an anomalous thematic vowel *a*: this was simply replaced by the standard *e/o*. This was also a proto-Italic phenomenon, as shown by the form *didet* ‘gives’ in Vestinian (a close relative of Oscan) in place of **didat* from the athematic reduplicated present **di-dh₃-ti* (Meiser 2003: 53); the same form underlies Latin third conjugation *reddit* ‘gives back’ < **re-didet*. This process will prove crucial to the development of heteroclis of the ‘*domō* type’, discussed in section 3.3.

One result of the strong trend in favour of thematic present stems pursued by the Italic branch is thus that the third conjugation brings together many verbs belonging to the oldest levels of the vocabulary and representing a large array of

⁶⁴ In the citation of Sabellic forms from inscriptions, bold face is conventionally used to transcribe words written in the native alphabet, which differs slightly from the Latin alphabet; words originally written in the Latin alphabet are given in italics.

different PIE present stem formations, most of which are no longer recognizable as distinct types. This further complicates the formal relationship between the present stem of these verbs and their other stems, a relationship made unpredictable anyway by the retention of both perfect and aorist formations for the perfect stem.

The same unpredictability characterizes the paradigm of verbs belonging to the Latin second conjugation (though to a lesser extent), owing to its own various origins. The two major lexical contributions to the second conjugation are made by the causative/iterative verbs in **-eye/o-* (e.g. *moneō* ‘warn’ < **mon-eye/o-*) and the statives in **-eh₁-*, perhaps extended to **-eh₁-ye/o-* by means of the productive thematic suffix (e.g. *rubeō* ‘blush/be red’ < **h₁rud^h-éh₁-*), which fell together formally at the proto-Italic stage largely thanks to the loss of intervocalic *y* (Meiser 1998: 189f.; Sihler 1995: 530-32). As these derived verbs did not possess any other stems in PIE, perfect and third stems are often created for them secondarily, in part by means of a process discussed in detail in section 3.3. This points to an important feature of the Latin verb which distinguishes it from that of PIE and marks out its further development away from the original derivational system: in Latin, it is the norm for each present stem to be accompanied by corresponding perfect and third stems, together making up the complete paradigm. However, on occasion we can see traces of the earlier system. For example, Latin second conjugation *pendeō* ‘hang’ is historically a derived essive built on the root **(s)pend-* ‘stretch’ (*LIV*²: 578) and as such it inherited no dedicated perfect and third stems; but rather than innovating new stems, its paradigm is completed by the corresponding stems proper to *pendō* ‘weigh’ (namely *pependī pēnsum*), which in PIE terms is its basic counterpart formed on the same root. Such instances constitute one factor leading to

unpredictability in the relationship between the stems in verbs of the second conjugation.

I return to this phenomenon too in section 3.3, as it is of particular relevance to the type of heteroclisis treated there, which I label the ‘*domō* type’. The fourth conjugation of Latin, meanwhile, is discussed in section 3.2 immediately below. I thus conclude this section by mentioning the major sources for the Latin first conjugation. There are two of these. A stem in *-ā-* < **-eh₂-* arose from the deadjectival formation exemplified above by **néw-e-h₂-* ‘renew’, from **néw-e/o-* ‘new’; meanwhile, denominal **-eh₂-ye/o-* produced thematic **-ā-ye/o-* (Meiser 1998: 186). These two also became identified at the proto-Italic stage, perhaps purely by sound change, and together explain the status of the first conjugation as the chief productive conjugation in Latin. As they were both derived formations in PIE terms, however, they brought with them no inherited perfect or third stems. The importance of this fact for ‘*domō*-type’ heteroclisis will, again, be treated in section 3.3 below.

3.2 The *capiō* type

3.2.1 Morphological properties of the *capiō* type

The following tables (3.3, 3.4, 3.5) present the inflectional behaviour of the first group of Latin heteroclite verbs to be discussed here, the type exemplified by *capiō* ‘take’, in comparison with the behaviour of the third and fourth conjugations, represented here by *regō* ‘rule’ and *sentio* ‘feel’ respectively. Only forms from the *inflectum* are given, as the distinctiveness of the *capiō* type lies entirely within this part of the paradigm: as was seen above, it is here that third and fourth conjugation forms are reliably opposed in terms of their morphology, and yet the *inflectum* of *capiō* is found to make use of forms patterning with both of these conjugations in different cells. For simplicity of presentation, out of the finite paradigm only the active portion is laid out here, since the finite passive forms of *capiō* show the same conjugational affiliations as their active counterparts.

As the juxtaposition of these three verbs shows, every inflected form of *capiō* patterns with the corresponding form in either the third or the fourth conjugation; no cell in its paradigm shows morphology which is independent from both and thus distinctive to *capiō* itself. To show this relationship more clearly, those parts of the paradigm of *capiō* which pattern with the third conjugation are underlined in Table 3.3, while those which pattern with the fourth are given in bold. This also brings out the distribution of the two types of inflectional behaviour over the *inflectum*. Strikingly, the generalization can be made that in all inflected forms where the stem vowel *-ī-* of the fourth conjugation is followed by another vowel – meaning that it appears as a short vowel on the surface – *capiō* patterns with the relevant fourth conjugation form; where it does not, *capiō* agrees with the third conjugation (cf. Rubenbauer and Hofmann 1995: 101).

Table 3.3 *Infectum* forms of *capiō* ‘take’

	pres. ind.	imperative	pres. subj.	future	impf. ind.	impf. subj.
1sg	<i>capiō</i>		<i>capiam</i>	<i>capiam</i>	<i>capiēbam</i>	<i>caperem</i>
2sg	<i>capis</i>	<i>cape</i>	<i>capiās</i>	<i>capiēs</i>	<i>capiēbās</i>	<i>caperēs</i>
3sg	<i>capit</i> ⁶⁵		<i>capiat</i>	<i>capiet</i>	<i>capiēbat</i>	<i>caperet</i>
1pl	<i>capimus</i>		<i>capiāmus</i>	<i>capiēmus</i>	<i>capiēbāmus</i>	<i>caperēmus</i>
2pl	<i>capitis</i>	<i>capite</i>	<i>capiātis</i>	<i>capiētis</i>	<i>capiēbātis</i>	<i>caperētis</i>
3pl	<i>capiunt</i>		<i>capiant</i>	<i>capient</i>	<i>capiēbant</i>	<i>caperent</i>
	present active infinitive		<i>capere</i>			
	present passive infinitive		<i>capī</i>			
	present participle		<i>capiēns</i>			
	gerund		<i>capiendum</i>			

Table 3.4 *Infectum* forms of (third conjugation) *regō* ‘rule’

	pres. ind.	imperative	pres. subj.	future	impf. ind.	impf. subj.
1sg	<i>regō</i>		<i>regam</i>	<i>regam</i>	<i>regēbam</i>	<i>regerem</i>
2sg	<i>regis</i>	<i>rege</i>	<i>regās</i>	<i>regēs</i>	<i>regēbās</i>	<i>regerēs</i>
3sg	<i>regit</i>		<i>regat</i>	<i>reget</i>	<i>regēbat</i>	<i>regeret</i>
1pl	<i>regimus</i>		<i>regāmus</i>	<i>regēmus</i>	<i>regēbāmus</i>	<i>regerēmus</i>
2pl	<i>regitis</i>	<i>regite</i>	<i>regātis</i>	<i>regētis</i>	<i>regēbātis</i>	<i>regerētis</i>
3pl	<i>regunt</i>		<i>regant</i>	<i>regent</i>	<i>regēbant</i>	<i>regerent</i>
	present active infinitive		<i>regere</i>			
	present passive infinitive		<i>regī</i>			
	present participle		<i>regēns</i>			
	gerund		<i>regendum</i>			

Table 3.5 *Infectum* forms of (fourth conjugation) *sentīō* ‘feel’

	pres. ind.	imperative	pres. subj.	future	impf. ind.	impf. subj.
1sg	<i>sentīō</i>		<i>sentiam</i>	<i>sentiam</i>	<i>sentīēbam</i>	<i>sentīrem</i>
2sg	<i>sentīs</i>	<i>sentī</i>	<i>sentīās</i>	<i>sentīēs</i>	<i>sentīēbās</i>	<i>sentīrēs</i>
3sg	<i>sentit</i>		<i>sentiat</i>	<i>sentiet</i>	<i>sentīēbat</i>	<i>sentīret</i>
1pl	<i>sentīmus</i>		<i>sentīāmus</i>	<i>sentīēmus</i>	<i>sentīēbāmus</i>	<i>sentīrēmus</i>
2pl	<i>sentītis</i>	<i>sentīte</i>	<i>sentīātis</i>	<i>sentīētis</i>	<i>sentīēbātis</i>	<i>sentīrētis</i>
3pl	<i>sentiant</i>		<i>sentiant</i>	<i>sentient</i>	<i>sentīēbant</i>	<i>sentīrent</i>
	present active infinitive		<i>sentīre</i>			
	present passive infinitive		<i>sentīrī</i>			
	present participle		<i>sentīēns</i>			
	gerund		<i>sentīendum</i>			

⁶⁵ Strictly speaking, there is no distinction between third and fourth conjugation behaviour in the form *capit*: cf. third conjugation *regit*, fourth conjugation *sentit*. I associate it with the third conjugation here on the strength of the corresponding passive *capitur*, cf. third conjugation *regitur* as opposed to fourth conjugation *sentitur*. The short *i* seen in the active *sentit* is in fact the result of a historical process of vowel shortening before final *t* and *m* (Meiser 1998: 77), which led to a transparent synchronic constraint ruling out long vowels in those positions (cf. the subjunctive forms *capiam*, *capiat* as against *capiās*). Short vowels before 3pl *-nt* have a different origin (Meiser 1998: 75).

The number of basic verb stems belonging to the *capiō* type is small in comparison with the numbers belonging to the third and fourth conjugations. A comprehensive list of the simplex verbs straightforwardly showing this heteroclite behaviour in Classical Latin, drawn from de Vaan 2008, comprises only *capiō* ‘take’ itself, *cupiō* ‘desire’, *faciō* ‘do’, *fodiō* ‘dig’, *fugiō* ‘flee’, *gradior* ‘step’, *iaciō* ‘throw’, *morior* ‘die’, *orior* ‘rise’, *pariō* ‘produce’, *patior* ‘suffer’, *potior* ‘acquire’, *quatiō* ‘shake’, *rabiō* ‘rage’, *rapiō* ‘snatch’, *sapiō* ‘taste (of)’, and *speciō* ‘watch’; more marginal members of the group are *sarpiō* ‘prune’ (which usually inflects as third conjugation *sarpō*), defective *aiō* ‘say’, archaic *horior* ‘encourage’, and *laciō* and **apiō*, which historically underlie existing prefixed verbs such as *ēliciō* ‘elicit’ and *coepiō* ‘begin’ respectively but are not found in their own right.⁶⁶ However, it will be noted that several of these verbs are central items of vocabulary, especially *faciō* ‘do’ and *capiō* itself; meanwhile, many prefixed derivatives sharing this inflectional behaviour are also found, some of which are more prominent in the language than the associated simplex verbs (e.g. *aggredior* ‘attack’ from *gradior* ‘step’). This type is thus more heavily represented in terms of textual frequency than in terms of its extension over the lexicon.⁶⁷

In synchronic and diachronic presentations of the Latin verb system, the *capiō* type is usually grouped with the third conjugation for purposes of exposition, though its divergence from the third conjugation proper is acknowledged by means of a label such as ‘3b’ (e.g. Meiser 1998: 194) or ‘third conjugation in *-iō*’ (e.g. Allen and Greenough 1903: §188, Weiss 2009: 399f.). This is largely for reasons of

⁶⁶ The simplex verb *laciō* is cited by the grammarian Festus (second century AD), but merely in order to provide a notional source of derivatives such as *ēliciō*; it is otherwise unattested, although it surely existed in that form at some point (cf. *laccessō* ‘provoke’ and the noun *laqueus* ‘snare’).

⁶⁷ Thus, for example, Denooz 2010 counts fourteen verbs of the *capiō* type among the hundred most frequently used verbs in the language; meanwhile, of the remainder, only thirteen belong to the first conjugation, the most populous and productive Latin conjugation.

tradition, as the Latin grammarians themselves identified conjugation membership principally on the basis of the 2sg present active indicative form (see e.g. Taylor 1991: 86-91) and therefore assigned verbs behaving like *capiō* to the third conjugation.

However, in line with the notion that the four regular conjugations are characterized in the *inflectum* by underlying stem vowels which are to be distinguished from the inflectional endings proper, but may influence their shape and interact with them (morpho)phonologically to a greater or lesser extent (cf. section 3.1.1), the same presentations often point out that on the whole it is possible to account for the *inflectum* forms of the *capiō* type simply by identifying in them a consistent underlying stem vowel *-i-* (e.g. Weiss 2009: 399f., Sihler 1995: 536f.), which differs from the stem vowels associated with the four main conjugations. Sihler states the synchronic situation as follows: ‘the difference between the 4th conj. and the *capiō* subtype of the 3rd conj. is that the stem-vowel is morphophonemically *-ī-* in the former and *-ĩ-* in the latter.’ For example, if the finite present indicative active forms of the fourth conjugation are taken to reflect the addition of the personal endings *-ō -s -t -mus -tis -unt*⁶⁸ to an underlying stem vowel *-ī-* (allowing for the operation of phonological principles reducing this to *i* before a following vowel or a final *t*, as in fnn. 54 and 65 above), the corresponding forms of the *capiō* type can all be derived trivially by assuming the addition of the same endings to a stem in underlying short *-i-* which undergoes no modification at all. For comparison, and to show how the forms of *capiō* come to coincide partially with those of the third conjugation as well as those of the fourth according to this view, I add a derivation of the corresponding forms of third conjugation *regō* in a similar spirit – assuming for

⁶⁸ In fact Sihler’s presentation uses slightly different underlying forms for these endings, but the difference is irrelevant here.

the time being that its stem vowel is an underlying *-e-* which surfaces as *i* before consonants but as zero before vowels.

- (3.1) *sentī- + -ō -s -t -mus -tis -unt > sentiō sentīs sentit sentīmus sentītis sentiunt
*kapi- + -ō -s -t -mus -tis -unt > capiō capis capit capimus capitis capiunt
*rege- + -ō -s -t -mus -tis -unt > regō regis regit regimus regitis regunt

In fact, as can be seen from Table 3.3 above, the majority of the *inflectum* forms of *capiō* do contain a short *-i-* immediately preceding the inflectional ending, which could be identified as this stem vowel. Its failure to appear in forms such as active infinitive *capere* (as opposed to **capire*) and 2sg imperative *cape* (as opposed to **capi*) only slightly compromises this generalization, as short *i* is almost never found in the relevant phonological contexts in Latin, thanks to the historical operation of specific sound changes, to which I return below (unstressed medial *i* > *e* before syllable-initial *r*, and final *i* > *e*: Meiser 1998: 68, 72). If these few forms without *i* on the surface can also be taken to reflect a stem vowel *-i-* at an underlying level, it emerges that the inflectional endings required to generate the *inflectum* forms of the *capiō* type are precisely those which can be taken to generate the regular third conjugation forms, given the necessary assumptions about the behaviour of the third conjugation stem vowel.⁶⁹

If membership of Latin conjugation classes is taken to depend on the identity of the stem vowel of the *inflectum*, then to derive the *inflectum* forms of the *capiō* type synchronically in this way is evidently tantamount to positing the existence of a fifth regular conjugation in Classical Latin, distinct from the four generally recognized explicitly and introduced above. On the other hand, if the crucial factor is taken to be the full complement of endings following this stem vowel, the traditional

⁶⁹ If the theme vowel is separated off from the inflectional endings proper in this way, it is possible to limit the necessity of positing different endings in the third and fourth conjugations to a single form, the present passive infinitive, which in the fourth conjugation takes the ending *-rī* (following the stem vowel) and in the third takes *-rī* (with no trace of the stem vowel). Here the *capiō* type patterns with the third conjugation.

identification of the type as a third conjugation variant can be retained, the divergence between the inflection of *capiō* and *regō* being based purely on the difference between their stem vowels. In fact, there is little genuine disagreement between these two conceptions. Notably, however, whatever the precise terms it is couched in, the analysis in terms of a uniform stem vowel *-i-* makes no substantial use of the notion of heteroclisis, or the distinction I have drawn above between third and fourth conjugation forms within the paradigm of *capiō*. On the assumption that three different underlying formations are reflected in each case, the resemblance between *capiō* and *sentio* (against *regō*), on one hand, and between *capis* and *regis* (against *sentis*), on the other, is of no particular morphological significance, and accordingly it would be misleading, or at least superficial, to identify heteroclisis here: despite appearances, there is no sense in which *capiō* is a fourth conjugation form, nor indeed is *capis* a third conjugation form in the sense intended in Table 3.3. Rather, all the resemblances noted above between inflectional formations in the *capiō* type and those in the third and fourth conjugations are to be explained as coincidences arising from the contextual neutralization of distinctions between underlying stem vowels, as illustrated in the derivations from **sentī-*, **kapi-* and **rege-* suggested above: in certain morphophonological contexts either the fourth conjugation stem vowel *ī* or the third conjugation stem vowel *e* may surface as *i*.

It is easy to understand the appeal of this analysis which makes no recourse to heteroclisis, as the prevalence of short *i* across the paradigm of *capiō*, whether in forms patterning with *regō* or with *sentio*, parallels that shown by the better-established stem vowels in the conjugations they characterize. In terms of formal neatness, this presentation of the facts clearly has much to recommend it, and as will be seen shortly, it in fact captures something of the diachronic background to the

Classical Latin state of affairs. However, this neatness comes at the important cost of relegating to the margins the defining synchronic property of the inflection of the *capiō* type, namely the fact that while its paradigm in the *inflectum* is clearly distinct from that of the four regular conjugations, this is not because it possesses any identifiable inflectional behaviour of its own. On the contrary, as shown above, every inflectional formation shown by this relatively small group of verbs is at least superficially shared with one or other of the (much more populous) third and fourth conjugations. This being the case, the psychological reality of an abstract synchronic analysis which posits that speakers nonetheless treated the inflection of *capiō* as underlyingly distinct from both the third and fourth conjugations in every form must surely be called into question.⁷⁰

Moreover, unlike the stem vowel *ī* of the fourth conjugation, which is also employed (for example) in the perfect and third stems of many fourth conjugation verbs (*-ī-vī -ī-tus* etc.), there is no evidence in support of a proposed stem vowel *i* from outside the *inflectum* forms of *capiō* which are at issue. Meanwhile, there are synchronic links between the *capiō* type and the fourth conjugation independently of the question whether these forms in *i* themselves represent fourth conjugation morphology, as I claim. Certain verbs which are listed above as belonging to the *capiō* type do not always distribute apparent third and fourth conjugation morphology as would be expected, instead tending to pattern with the fourth conjugation outside the zones given in bold face in Table 3.3: for example *potior* ‘acquire’ often shows 3sg present indicative *potītur* and infinitive *potīrī* (as opposed to *potītur*, *potī*), while *orior* ‘arise’, whose 3sg present indicative is *orītur* as

⁷⁰ While Weiss (2009: 399) endorses the notion of an underlying stem /kapi-/, showing a stem vowel *i* particular to this type, it is telling that he refers to Latin as possessing ‘four and a half conjugations’, explicitly commenting in a footnote ‘We say four and a half conjugations because the 3rd-*iō* shares all of its forms either with the 3rd or the 4th conjugations’ (fn. 1). He later comments (p. 400) that ‘the most efficient synchronic analysis of these conjugations [3rd and 3rd-*iō*] is not obvious’.

expected for the *capiō* type, only ever makes use of the infinitive *orīrī* (Martzloff 2006: 276). Such forms in long *ī* belong unambiguously with the fourth conjugation, as they cannot be derived from a putative stem in short *i*; their existence makes it hard to argue that speakers failed to associate other forms of the same verbs, such as 1sg *orior* and *potior* themselves, with the fourth conjugation as well. Furthermore, compounded forms of both *orior* and *pariō* ‘produce’, such as *adorior* ‘attack’ and *comperiō* ‘discover’, inflect as regular fourth conjugation verbs throughout.

In light of these observations it seems to do more justice to the synchronic status of the *capiō* type to identify it as heteroclite rather than a fifth Latin conjugation or a subtype of the third. Nonetheless, we will see shortly that this state of affairs is a fairly recent one in Classical Latin, and must have arisen within the attested history of the language; in diachronic terms, in fact, the *i*-stem label is highly appropriate for the *capiō* type, which was indeed independent from both the (ancestral) third and fourth conjugations at an earlier stage. But the deeper history of the *capiō* type is more obscure. As seen in the following section, an investigation of this history raises questions concerning both the morphological formations inherited from PIE by Italic and the original rationale behind the distribution of conjugation class membership over the lexicon.

3.2.2 *The origins of heteroclisis in the capiō type*

The investigation of heteroclite nominal behaviour in Asia Minor Greek carried out by Karatsareas 2011a, and discussed in section 2.6.1 above, demonstrates one way in which heteroclisis may arise in the paradigm of a lexeme: namely, in certain cells the paradigm may take on innovative morphological behaviour which previously characterized a different inflectional class. In the AMGk case, the innovative behaviour taken on is that of the *i*-stem declension originally restricted to neuter

nouns; the fact that in heteroclite paradigms the *i*-stem morphology represents the innovation, and not the earlier state of affairs, is evident both because it has affected several different existing declensions and thanks to historical and comparative evidence from elsewhere in Greek. In principle, a similar diachronic scenario might be envisaged for the Latin heteroclite type under discussion here: the central question concerning the origin of the type would then be whether verbs such as *capiō* belonged originally to the third conjugation before borrowing inflectional behaviour from the fourth in certain cells, or the reverse, incorporating third conjugation morphology into a previously fourth conjugation paradigm. The recorded history of the type within Latin would be of no help in distinguishing between these two potential accounts, as verbs of the *capiō* type show affiliation with both the third and fourth conjugations from the time of their earliest attestations.

However, evidence regarding the origins of this heteroclite type in Latin can be drawn from its sister languages in the Sabellic branch of Italic, and the upshot is to suggest that the truth of the matter lies with neither of these accounts. It was noted above that the Sabellic varieties possess counterparts of each of the four conjugations found in Latin: Rix (1994: 28) indeed points out that this is one of the principal arguments in favour of an Italic family within Indo-European, comprising both Sabellic and Latino-Faliscan. In particular, Sabellic verb forms are known which correspond to Latin fourth conjugation forms, with a long \bar{i} preceding the inflectional ending: for example Umbrian **persnihmu** (*persnihimu*, etc.) ‘pray!’ < **persnī-mōd* (Martzloff 2006: 225), in which the commonly used spellings **ih**, *ihi* reflect the length of the vowel, and Umbrian **seritu** ‘conserve!’, plausibly representing **serwī-tōd* (from a fourth conjugation equivalent of Latin first conjugation *servō*: Meiser 1998: 194). But crucially, alongside these, Sabellic also shows evidence of a class of

verbs which (at least historically) employed a short *i* in the corresponding position. One member of this class in Oscan is the verb ‘do’, which clearly represents the same etymon as Latin *faciō*. The stem vowel can be seen in the 3sg present subjunctive form **fakiiad** (cf. Latin *faciat*), in which the second **i** marks the automatic glide occurring between the two vowels in hiatus (Buck 1904: 28); that it is short is demonstrated by the fact that it has undergone syncope, as is regular in Oscan for short vowels in certain medial contexts, in the future imperative form *factud* < **faki-tōd* (where, meanwhile, its presence at an earlier historical stage is established by the fact that the regular development of inherited **kt* to *ht* has not taken place, cf. **ehtrad** ‘outside’ < **ektrād*, Meiser 1986: 92). The existence of this short *i* formation in Umbrian as well is demonstrated by the relationship between infinitive **herifi** ‘want’ and 3sg present passive **herter** ‘it is wanted (i.e. one should)’ < **heri-tir* (Meiser 1998: 194), where syncope of the *i* has likewise taken place.

As has long been recognized (cf. already von Planta 1892: 248-58 and the literature cited there), this Sabellic class clearly invites comparison with the *capiō* type in Latin, and should have consequences for our understanding of its history. The existence of a contrast common to the Sabellic languages between verbs bearing historic long and short *i*, identifiable by means of evidence internal to Sabellic, makes it especially difficult to imagine that the prevalence of short *i* in the inflection of the *capiō* type comes as the coincidental result of the adoption of third conjugation morphology into a fourth conjugation paradigm, or vice versa. Instead, the short *i* seen in both (synchronically fourth conjugation) *capiō* and (synchronically third conjugation) *capis* is likely to reflect a single stem formant inherited by both branches of Italic. That is to say, Sihler’s derivation of the forms of *capiō* from a single stem distinct from both third and fourth conjugations should after all be

endorsed, but recast as a *diachronic* statement: whatever its synchronic status may be, the *capiō* type represents an Italic formation distinct from those which are continued as the third and fourth conjugations in both Latin and Sabellic.

This being the case, the origin of heteroclisis in verbs of the *capiō* type must be quite different from that seen in the case of the AMGr nouns discussed above, because the synchronic status of the *capiō* type as heteroclitite in Classical Latin relies not on the innovative adoption of morphology from elsewhere in the language, but on the loss of the separate morphological identity which it previously possessed, by which it stood in contrast with both the third and fourth conjugations. How did these two contrasts come to be erased to give the Classical Latin state of affairs?

A treatment of the complicated diachronic relationship between the morphology of the *capiō* type and that of the Latin fourth conjugation belongs in the context of a discussion of the ultimate historical connections of the Italic short *i* stem formation, to which I turn shortly; although the questions involved are far from settled, it will emerge that where the *capiō* type patterns with the fourth conjugation in Classical Latin, the loss of whatever earlier formal distinctions previously existed between the two is likely to date to an early stage in the prehistory of Latin.

Meanwhile, however, the loss of the *capiō* type's distinctive identity vis-à-vis the *third* conjugation in fact rests on a very general and well-established development which not only left an abundance of clear traces in the structure of Latin, but which emerged within the recorded history of the language and can thus be tracked directly in our attestations: this is the phenomenon (more precisely, the array of related phenomena) commonly known as 'vowel weakening', which operated from around the fifth century to around the third century BC. Vowel weakening had the overall effect of reducing the number of distinct short vowels available in non-

initial syllables, and is generally taken to be the result of a general assignment of stress to the initial syllable in the Latin of the archaic period;⁷¹ in certain contexts the eventual result could be syncope, the loss of the vowel in question altogether. For our purposes here, the crucial sound changes involved are as follows (for more detailed exemplification see e.g. Meiser 1998: 66-74, Leumann 1977: 79-95). In absolute final position inherited *i* and *e* fell together to give *e* (Latin *ante* ‘before’, cf. Ancient Greek *antí* ‘against’), while in a final syllable before a single consonant they fell together as *i* (cf. the third declension genitive ending *-es* still seen in third-century epigraphic *salutes* ‘of Safety’ (*CIL* I²: 450), corresponding to Classical Latin *salutis*); an analogous raising process caused the attested changes of archaic nominative singular *-os* to *-us* and archaic 3pl *-ont* to *-unt*. Meanwhile, in a medial open syllable the distinctions between all five vowels are neutralized. The regular result in most cases is *i*, as seen in e.g. *angina* ‘throat infection’, the outcome of borrowed Ancient Greek *ankhónē* ‘strangling’, as well as in alternations such as *premo* ‘press’ – *comprimō* ‘press together’; however, before *r* (including *r* resulting from the rhotacism of intervocalic *s*) the regular result is *e*, as demonstrated in the name of the town *Faler-iū* in comparison with that of its inhabitants the *Falis-cī*.

Recall (cf. section 3.1.3) that the third conjugation of Italic principally continues the ‘simple thematic’ stems of Indo-European, which were characterized in the present indicative by a short vowel appearing in the guises **-e/o-* before the personal endings. Except in the 1sg form *-ō*, which reflects an earlier combination of thematic **o* with a following laryngeal, this short vowel survived into Italic, although the distribution of **e* and **o* over the persons of the paradigm is sometimes subject

⁷¹ A monographic treatment of this topic is Nishimura 2008, which also treats the question (not discussed in the present work) whether syncope appeared as an alternative to weakening proper or as the ultimate stage in a weakening process. See his p. 9, fn. 19 for a comprehensive bibliography on the notion of initial stress in Latin and the other Italic languages.

to analogical change (cf. Marrucinian 3pl *feret* ‘they bear’, as if **b^hér-e-nti* rather than inherited **b^hér-o-nti* > Latin *ferunt*: Untermann 2000: 275). However, it would naturally have been affected by the processes of vowel weakening outlined above, as indeed 3pl *ferunt* shows. Seen in this light, the *i* appearing before the personal endings in the remaining forms of the present indicative of the third conjugation (2sg *-is*, 3sg *-it*, 1pl *-imus*, 2pl *-itis*) can be identified as the *expected* outcome of the inherited short vowels after vowel weakening, continuing earlier **-e-s*, **-e-t*, **-o-mos*, **-e-tes*; the same goes for the 2pl imperative *-ite* < **-e-te*. In these forms, the third conjugation would thus have come to resemble the *capiō* type without undergoing any disruption at all in morphological terms. But at the same time, in the 2sg imperative, but also in the active infinitive and throughout the imperfect subjunctive, a short *i* in the forms of the *capiō* type would automatically have undergone lowering to *e* as a result of its phonetic environment (in absolute final position or before *r*). Here, then, it was the *capiō* type that came to pattern with the third conjugation, rather than the reverse; again, however, the upshot was to erase the distinction between the morphological behaviour of the two classes of verbs in the relevant cells.

(3.2)

<i>*reg-e-s</i>	>	<i>regis</i>	<i>*cap-i-s</i>	>	<i>capis</i>
<i>*reg-e-t</i>	>	<i>regit</i>	<i>*cap-i-t</i>	>	<i>capit</i>
<i>*reg-o-mos</i>	>	<i>regimus</i>	<i>*cap-i-mos</i>	>	<i>capimus</i>
<i>*reg-e-tes</i>	>	<i>regitis</i>	<i>*cap-i-tes</i>	>	<i>capitis</i>
<i>*reg-e-te</i>	>	<i>regite</i>	<i>*cap-i-te</i>	>	<i>capite</i>
<i>*reg-e</i>	>	<i>rege</i>	<i>*cap-i</i>	>	<i>cape</i>
<i>*reg-e-re</i>	>	<i>regere</i>	<i>*cap-i-re</i>	>	<i>capere</i>

In this way, every form of *capiō* which did not already pattern with the fourth conjugation came to share the behaviour of the third. The loss of a distinctive morphological identity for the *capiō* type, yielding to heteroclisis between two independently existing conjugations, thus does not originate in developments

restricted to the morphology of the verb, or to the morphology at all; instead it rests on the operation of the diachronic *phonological* phenomenon of vowel weakening.

3.2.3 The prehistory of the *capiō* type and its relationship with the fourth conjugation

While our knowledge of the history of Latin enables us to understand how the *capiō* type would have come to pattern with the third conjugation in part of its paradigm given the existence up to that point of a consistent short *i* in the relevant cells, the diachronic identity of this short *i* remains to be discussed, as does the relationship between the *capiō* type and the fourth conjugation: as will be seen, these two questions are connected.

I will begin with the ground held in common by practically all who have considered the topics at issue here. As discussed in section 3.1.2 above, the verb system of PIE made use of a thematic suffix **-ye/o-* which derived present stems from the stems of nouns. In Italic, as in practically all IE branches, this formation was highly productive, and a major part of the productivity of the Latin first conjugation is due to the fact that it continues (among other things) the denominal formation in **-eh₂-ye/o-* which in the proto-language was built on the basis of stems in both **-eh₂-* and thematic **-e/o-* (> the Latin first and second declensions respectively); in Latin itself, the first conjugation may produce denominal verbs on the basis of stems of other types as well (e.g. *iudicō* ‘judge’ from *iudex* (genitive *iudicis*), *nōminō* ‘name’ from *nōmen* (genitive *nōminis*), Leumann 1977: 545). However, the fourth conjugation of Latin also contains numerous denominal verbs, which predominantly correspond to existing stems in *-i-*. Examples include *sitiō -īre* ‘be thirsty’ cf. *siti-s* ‘thirst’, *vestiō -īre* ‘dress’ cf. *vesti-s* ‘clothing’, *fīniō -īre* ‘end’ cf. *fīni-s* ‘boundary’ (Thurneysen 1879: 3-10); at the same time, fourth conjugation denominals based on nominal stems of other kinds also exist, e.g. *serviō -īre* ‘serve’

cf. *servus*. The Umbrian fourth conjugation verb giving the form **persnihimu** ‘pray!’ noted above is also to be taken as a denominal, from a noun **perks-ni-* ‘devotional act’ related to Latin *prex* ‘prayer’ (Martzloff 2006: 691). This pattern, taken together with comparative evidence from other IE languages (e.g. Ancient Greek *mēniō* ‘be angry’ cf. *mēni-s* ‘rage’, shows that one of the historical components underlying the Latin fourth conjugation is an inherited thematic denominal formation in **-i-ye/o-*.

At the same time, though, the Latin fourth conjugation also contains many basic verbs, including *sentīō -īre* itself, suggesting that this denominal formation is not its only PIE source. Although this conclusion has occasionally been doubted (Graur 1937), comparative evidence seems to demonstrate that the fourth conjugation can also continue the productive late IE thematic formation in which the verb root is followed by **-ye/o-*, in cases where the reflex of this formant comes to follow a consonant in Italic: for example, *sepeliō -īre* ‘bury’ corresponds to Sanskrit (3sg) *saparyāti* ‘honour’ < **sepel-ye/o-*, and *veniō -īre* ‘come’, like Ancient Greek *baínō*, reflects **g^wm-ye/o-* (Meiser 1998: 195). If this is the case, it entails that at some point in their development the distinction between the **-C-i-ye/o-* and **-C-ye/o-* seen in these two formations was lost, the result being the morphology of the fourth conjugation.

This much is commonly accepted. However, the existence of the *capiō* type complicates this scenario. This is because comparative evidence is also found that links verbs of the *capiō* type with the same PIE suffix **-ye/o-* which has just been identified as giving rise to Italic fourth conjugation morphology. For instance, the parallelism between *speciō* ‘watch’ and, with the same meaning, Sanskrit (3sg) *pásyati* and Ancient Greek *sképtomai* (<**sképyomai*, showing consonantal metathesis) suggests a common etymon **spék-ye/o-*; the correspondence between

cupiō ‘desire’ and Sanskrit *kupyati* ‘be (emotionally) moved’ likewise suggests thematic origins for the *cupiō* type (cf. *LIV*²: 575f.; 359 respectively). The question arises, then, whether both the fourth conjugation and the *cupiō* type, which we have established as separate verb classes within Italic, nonetheless ultimately continue a single PIE formation in **-ye/o-* – and if so, how they came to diverge. The alternative is that the parallels provided should be taken as the result of independent developments in the languages involved, and that the *cupiō* type in fact developed from a somewhat different starting point in PIE.

This alternative view has a long history, most of its variants drawing on the notion of a ‘semi-thematic’ pattern (inherited by certain branches of IE only) whose classic formulation appears in Meillet (1908: 109-14). According to this proposal, rather than continuing a present stem in **-ye/o-* in which the thematic vowel originally appeared throughout but alternated according to person, the *cupiō* type would reflect an inherited alternation between (thematic) **-yo-* and (athematic) **-i-*, the latter simply representing the same suffix seen in the former (and in normal **-ye/o-* itself) but without the thematic vowel. Scholars have taken this putative pattern, incorporating athematic inflectional morphology, to explain a range of different peculiarities found in the verb systems of the IE family, with no general consensus on the precise branches to which the phenomenon should be supposed to extend or the attested phenomena which it might explain; for an outline of the uses to which the semi-thematic conjugation has been put, see Jasanoff (1978: 96f). In the Italic case, the obvious benefit of this reconstruction would be to provide a direct way to account for forms such as 2sg *capis*, whose agreement in inflection with historically thematic *regis* is, as we have seen, secondary: if *capis capit* simply reflects an inherited athematic formation in **-i-si* **-i-ti* (alongside thematic *cupiō*,

capiunt < *-yo-h₂, *-yo-nti), no special explanation is required for the forms we find, or for the distinct existence of a short *i* class in Italic.⁷²

In recent times, the potential of this idea to disentangle the history of the *capiō* type and the fourth conjugation has been championed most prominently by Schrijver 2003, and I will take his position (followed e.g. by de Vaan 2011) as representative of this side of the debate: on the grounds that it is impossible to explain both the *capiō* type and the fourth conjugation as continuing the same PIE formation in *-ye/o-, he sees in *capis*, *capit* a demonstration of the reality of inherited athematic *-i-. In fact, Schrijver does not endorse the idea of a semi-thematic pattern lying behind the *capiō* type, as he points out (p. 59) that Latin tends to generalize thematic morphology beyond its original bounds anyway, meaning that the apparently thematic inflection of *capiō*, *capiunt* could be secondary and does not need to be projected back to the PIE stage: he prefers to reconstruct a *fully athematic* paradigm with a stem in *-i- (earlier *-i-mi *-i-si *-i-ti etc.), which he also identifies as surviving in Balto-Slavic (following Kortlandt 1989: 109).

Schrijver's argumentation takes as its starting point his reconstruction of the way in which basic verbs in *-ye/o- came to fall together with the denominals in *-iye/o-, the resulting type giving rise ultimately to the Latin fourth conjugation as it is attested. As he points out, a sound change which would produce this result, in which post-consonantal *y* becomes syllabic before a vowel, is independently reconstructed for the history of Latin. Its effects can be seen in (trisyllabic) *medius* 'middle (adj.)', in comparison with (disyllabic) Sanskrit *mādhyas* and Gothic *midjis*,

⁷² It is true that the reconstruction of a semi-thematic pattern lying behind the *capiō* type would still require some kind of explanation for 1pl *-imus*, which does not obviously continue expected thematic *-yo-mos or the like; however, it seems reasonable to see here the effects of levelling, cf. *-is*, *-it*, *-itis* (whether before or after the vowel weakening process that would affect any medial *-yo-). But whatever the explanation, something similar is already required to account for fourth conjugation *-imus* from a starting point *-iyo-mos. In any case, in Schrijver's fully athematic account this issue does not arise.

all reflecting **med^hyos*. Similarly, Latin *maius* /majjus/ and Oscan *mais* < **mag-yos* ‘bigger’, which both display the result of a proto-Italic assimilation **gy* > *yy*, show that the inherited comparative suffix was **-yos* in proto-Italic; in Latin, however, this has become disyllabic after a consonant (giving trisyllabic *melius* ‘better’ etc.) thanks to the same sound change. For Schrijver (pp. 61f.), the loss of the distinction between verbs in **-ye/o-* and verbs in **-iye/o-* can be ascribed straightforwardly to this development, which would therefore give a single conjugation in **-C-iye/o-*. This formation is transparently found in both (basic) *sentio*, *sentiunt*, and (derived) *sitio*, *sitiunt*. Meanwhile, on this view, the *-ī-* of the remaining personal forms would emerge regularly from earlier **-i(y)o-*, **-i(y)e-* after vowel weakening: Schrijver compares forms such as *tībīcen* ‘flautist’ < **tībio-kan-* and the vocative in *-ī* (for expected *-ie*) of nouns in *-ius* (e.g. *filī* ‘O son’; this contraction is therefore termed the ‘*filī* rule’ in Weiss 2009: 122).

According to this conception, the creation of the Latin fourth conjugation from its two thematic sources took place thanks to sound changes operating not at the proto-Italic stage but within the history of Latin specifically. Schrijver takes this late date for the development to be unavoidable, on the basis that a late date is required in any case in order to explain the precise form of the fourth conjugation verb *sepeliō* ‘bury’. Pointing out that by ‘Exon’s Law’ (Sihler 1995: 70, following Cowgill 1970: 149, fn. 33),⁷³ a short vowel in the second syllable of a four-syllable (or longer) Latin word regularly underwent syncope if this syllable was open (e.g. *maximus* ‘biggest’ < **mag-isemo-*), Schrijver claims that the absence of syncope in *sepelīre* shows that the second syllable must still have been closed at the time this change operated, meaning that inherited **sepel-ye/o-* had survived up to this point.

⁷³ In fact, Schrijver’s statement of this ‘law’ is somewhat different from Sihler’s and Cowgill’s, which in turn do not agree with Exon’s (1906: 128) original observation. On this point see Sen (2012, especially pp. 205-7), who distinguishes between several ‘Exon’s Law syncope’s’.

By extension, for Schrijver, any verb inherited in the form **C-ye/o-* must have survived as such until the period of vowel weakening and syncope, an advanced stage in the history of Latin which postdates our earliest attestations of the language. But, crucially, this conclusion has the further consequence that it leaves no way to account for the forms of the *capiō* type on the basis of the same thematic formation: if **sepel-ye-s(i)* survived to a relatively late date and then developed to *sepelīs* entirely according to independently recognizable Latin sound changes, there is no principled way to explain how a notional **kap-ye-s(i)* could have split off at any point to undergo the separate development to *capis*. Schrijver (pp. 64f.) argues against several proposals purporting to recognize a change **yV > i* in individual instances (e.g. the notion that the third declension genitive *-is* reflects an earlier **-yes/yos*, Sihler 1995: 316f.); but strictly these rebuttals are unnecessary, as any such change could not in any case have affected verbs in **-ye/o-* without resulting in an open second syllable in the reflex of **sepel-ye/o-*. Ruling out the possibility that a thematic formation can explain the forms of the *capiō* type, Schrijver thus identifies it instead as based on athematic **i*, meaning that the relationship – if any exists – between the fourth conjugation and the *capiō* type belongs beyond comparative reconstruction and in internal reconstruction of the proto-language.

Schrijver finds (p. 60) that ‘the Italic evidence allows a straightforward conclusion’ in favour of this identification. He supports it with arguments from Celtic, suggesting that the Old Irish ‘BII presents’ (which also show traces of short *i* before the personal endings in some forms) partially continue both thematic **-ye/o-* and athematic **-i-* inflections, although he admits that ‘the evidence for athematic *i*-presents in Celtic is less clear-cut than in Italic’ (p. 71). However, his view that an athematic formation is required to explain the behaviour of the Old Irish BII class

even in part is not generally shared, as Schumacher (2004: 37-41) shows that the contribution of thematic **-ye/o-* (which is acknowledged by Schrijver in any case) is sufficient on its own. In fact, it is hard to find any independent support for a present in **-i-* separate from **-ye/o-*, as the many phenomena it has been used to explain can generally be dealt with in terms of better-established PIE or language-specific behaviour (see e.g. Nagy 1970: 20-26, 35 on Celtic and Germanic, signalled by Martzloff 2006: 286, 289, and Jasanoff 2008: 356 on Balto-Slavic). The burden of proof therefore lies on Schrijver to demonstrate that the reconstruction of an athematic **i* formation really is necessary in the Italic case.

Seen in this light, Schrijver's arguments are not sufficiently compelling. The crucial piece of chronological evidence he provides is the failure of *sepelire* to undergo syncope, which in his eyes must constrain the date at which the **-ye/o-* and **-iye/o-* verbs were conflated in Latin. However, as pointed out by Martzloff (2006: 288), syncope cannot be relied on to take place even in words which certainly fit the profile of (any suggested variant of) 'Exon's Law', such as *calamitās* 'damage' (which fails to give the phonologically acceptable **calmitās*). What is more, forms of the verb such as 2sg *sepelīs* or 3sg *sepelīt* (before final *-īt > -it*) would not have been subject to syncope anyway under such a 'law'; and Schrijver has no reason to suppose that these forms did not already exist at the time, except for the fact that the development of this long *ī* must come after the change **sepely-* > **sepeliy-* whose date is at issue in the first place.

Once 'Exon's Law' is recognized as a red herring for the question of chronology, it becomes possible to propose a far earlier date for the emergence of the fourth conjugation from its two sources. Note that in Schrijver's account, it goes unnoticed that Sabellic also has a class in long *ī* corresponding to that seen in Latin

(and likewise opposed to a short *i* class): this is because his chronology for **sepely-* > **sepeliy-* means that for him the developments in the two Italic branches must be historically independent. But given this striking parallelism, it is far more likely that both inherited the \bar{i} class in common from proto-Italic. The conflation of primary verbs such as **sepelye/o-* and denominals such as **sitiye/o-* into a single conjugation is most plausibly dated to this stage.

If the change **sepelye/o-* > **sepeliye/o-* took place in proto-Italic, the question we face is the following: does the failure of verbs such as *capiō* to undergo the same development (and thus their failure to fall together with the fourth conjugation ancestrally) mean that they did not inherit the sequence *-*Cye/o-* in the first place, as Schrijver suggests, or did they inherit the sequence but fail to undergo the same sound change? The latter answer requires us to identify some principled difference between the groups of verbs in *-*Cye/o-* which were affected by the change, like *sepeliō*, and those, like *capiō*, which were not.

But in fact just such a difference can be observed in the Italic verbs. As has often been observed and is laid out most comprehensively by Martzloff (2006: 275), practically all verbs of the *capiō* type share the property that their root contains a short vowel followed by a single stop (cf. *rapiō*, *cupiō*, *quatiō*, *patior*, *fodiō*, *gradior*, *faciō*, *iaciō*, *speciō* etc.): this emerges clearly from the list given above on p. 152. Meanwhile, underived fourth conjugation verbs (i.e. those not reflecting the denominal formation) tend to have heavy or polysyllabic roots (*dormiō*, *hauriō*, *audiō*, *sentiō*, *vinciō*; *sepeliō*, *amiciō*,⁷⁴ and cf. also Umbrian **amparitu**, **purtuvitu**); if this is not the case, the root ends in a sonant (*saliō*, *veniō*, *feriō*).

⁷⁴ Note that *amiciō* ‘clothe’, belonging to the fourth conjugation, is historically a prefixed form of *iaciō* ‘throw’ (< **amb-jak-* ‘throw around’), suggesting that the presence of the extra material affected this verb’s further development: this makes sense in the light of the general distribution noted here. However, transparently prefixed forms do not by any means reflect this generalization, as the

This distribution cannot be taken as phonologically motivated in synchrony (which would obviously make it unsuitable as an indication of the historical distribution of the two inflectional types, and indeed would cast doubt on the need to treat them as morphologically distinct types at all). For one thing, sonant roots may be found in the *capiō* type, notably *morior* ‘die’, *pariō* ‘produce’, while we have seen that *orior* ‘arise’, which to some extent makes up a semantic cluster with the two preceding verbs, also shows some third conjugation morphology; furthermore, of course, denominals such as *sitiō* < *sitis* ‘thirst’ belong to the fourth conjugation regardless of their shape. And yet the generalization in terms of root shape is strikingly accurate, with regard to the conjugational behaviour of underived verbs. It surely seems implausible, however, that verbs belonging to two separate morphological types, characterized by thematic **-ye/o-* and the proposed athematic **-i-*, would have happened to survive into Italic in such a way as to give rise to this observed distinction as a matter of lexical inheritance. Rather, the pattern makes sense as the result of differential treatment of what was originally a single formation. At the proto-Italic stage, on the basis of the shape of the root (Martzloff 2006: 282 following Niedermann 1908) or its prosodic weight (Weiss 1996), sound change came to distinguish **kapye/o-* and innovative **sepeliye/o-*, the latter now patterning with the denominals in **-iye/o-* to give the fourth conjugation as we know it. The hypothesis of a separate IE formation in athematic **-i-* is unnecessary, and cannot explain the pattern of conjugational behaviour observed over the lexicon.

The crucial point emerging from the foregoing discussion is thus the following: the same IE formation in **-ye/o-* is continued in Latin in both the *capiō* type and in the fourth conjugation, as (naturally) in their Sabellic equivalents. This

conjugational behaviour of the simplex verb influences that of the compound. It is notable, however, that *ad-orior* ‘attack’ and *ex-perior* ‘test’ show fourth conjugation morphology much more reliably than their simplex counterparts *orior*, *pariō* (Meiser 1998: 195).

means that the characteristic quantitative distinction, already present in proto-Italic, seen between the two in those parts of the paradigm where they differ, e.g. 2sg present indicative forms *capis* and *sepelīs*, must emerge from the distinction between **kapyes(i)* and **sepeliyes(i)*: the opposition between a conjugation in short *i* and a conjugation in long *ī*, which historically underlies e.g. the infinitive, was abstracted from pairs such as this. An examination of the details of the developments involved here would take us too far afield, as more than one sequence of events could lead to the distinction observed;⁷⁵ evidently, though, *-is* < **-ye-s(i)* (rather than *-is* < **-i-s(i)*, with Schrijver) is hardly implausible in view of the universally agreed development *-īs* < **-iyes(i)*.

At the same time, however, where the *capiō* type and the fourth conjugation pattern *together* in Latin, we are already in a position to understand why this should be: as observed above, Latin shows evidence of a sound change in which post-consonantal **y* became syllabic, giving trisyllabic *medius* < **med^hyos*. The agreement of *capiō* < **kapyō* with *sepeliō* < **sepeliyō* follows as a matter of course. The situation identified above for the *capiō* type at the stage before vowel weakening, in which it showed agreement only with the fourth conjugation, and only in those forms not possessing long *ī*, is thus explained.

3.2.4 Conclusion

The systematic behaviour of the Latin *capiō* type provides an example of heteroclisis very different both from lexically idiosyncratic instances, such as that seen in the declension of *domus*, and from the more wide-ranging and influential phenomenon

⁷⁵ The essential question is whether **-yes(i)* > *-is* occurred as the regular result of syncope, perhaps the very same phenomenon as that seen in **-iyes(i)* > *-īs* (Weiss 1996: 673), or whether, by contrast, a sound change **-iyes(i)* > **-iyis(i)* first gave rise to **-yes(i)* > *-yis(i)* by analogy (Martzloff 2006: 306f.). There are plausible arguments in both directions. In the latter case, however, it is not clear to me why the same analogy would not also have affected the ancestors of the other conjugations (e.g. a change **-āyes(i)* > **-āyis(i)* did not take place).

investigated by Karatsareas 2011a. While heteroclisis in the latter case came as the result of the adoption of morphology which already characterized a different inflectional class, in the *capiō* type we see a case where heteroclisis has arisen thanks to the *effacement*, purely by means of regular sound change, of previously existing morphological distinctions, which had separated the type from both the (ancestral) third and fourth conjugations in proto-Italic and prehistoric Latin. However, the lexical distribution of this heteroclite behaviour suggests that this isolation was itself not original and does not mark out the *capiō* type as continuing a special athematic IE stem formant; instead, the *capiō* type shares its origins with part of the fourth conjugation, continuing the productive thematic **-ye/o-* of late IE which underpins so much of the Italic verb system.

Finally, it is worth observing, in the light of Maiden's proposal for a constraint on the synchronic behaviour of heteroclisis, that it is not clear that it is appropriate to say that the distribution of conjugational affiliation shown by the *capiō* type follows a pre-existing distribution in stem alternants over the paradigm (or, here, the portion of the paradigm affected). It is true that, as noted at the start of the chapter, the cells characterized by fourth conjugation morphology in *capiō* are those in which we see a short *i* rather than a long *ī* in fourth conjugation verbs. However, in synchronic terms, this can hardly be said to represent a morphological alternation in the fourth conjugation paradigm, as a general phonological constraint rules out long vowels in prevocalic position (cf. fn. 54). In fact, it is unlikely that these parts of the fourth conjugation paradigm ever contained long vowels: we have seen that the paradigm of *sepiō* continues **-i-ye/o-* directly. All this is to skirt the obvious point that there is no reason to expect heteroclisis which arises purely by means of sound change, as here, to conform to pre-existing morphological patterns:

rather, of course, sound change is one means by which such morphological patterns can appear in a language.

3.3 The *domō* type

In this section and the next, I treat two further instances of heteroclisis which can be recognized in the Latin verb system, referred to here as the *domō* and *petō* types. Synchronically these belong together, in opposition to the *capiō* type, in an important respect. While the paradigm of *capiō* is identifiable as heteroclite owing to the unexpected morphological relationship which it shows between different forms within the *inflectum*, this is not what marks out *domō* ‘tame’ or *petō* ‘seek’: instead, verbs of these types show an unexpected contrast in conjugational affiliation between the *inflectum* and the rest of the paradigm (i.e. the perfect and third stem forms). As these sections will show, however, the two heteroclite groups differ greatly with regard to the diachronic developments which underlie their behaviour.

Given the claim, made in section 3.1.1 above, that four conjugations are recognized for Latin on the basis of distinctions found within the morphology of the *inflectum*, it may seem paradoxical to attempt to point to heteroclisis involving a conjugational opposition between the *inflectum* and any other set of cells. However, although the present, perfect and third stems do show a great deal of independence in at least some parts of the Latin verb system, some salient regularities dependent on conjugation class do nonetheless obtain across the paradigm.⁷⁶ The *domō* and *petō* types stand out as heteroclite by the way in which their morphology cross-cuts these regularities.

The applicability of the term in such cases emerges clearly from a presentation of the behaviour of the verbs involved. In Table 3.6 below, I provide a list of the simplex verbs belonging to the *domō* type in Classical Latin, giving

⁷⁶ See e.g. Aronoff (1994: 49).

representative forms showing the present, perfect and the third stems (the present active infinitive, 1sg perfect active indicative and supine respectively).

The unusual behaviour of these verbs, against the background of the Latin verb system as a whole, is easily recognized on the basis of the forms given in this table. The first conjugation, showing stem vowel *-ā-* in the *infectum*, is both the most populous and the most productive conjugation in Latin, and the behaviour of the inflected forms of verbs which make use of this element in the *infectum* is overwhelmingly predictable over the rest of the paradigm: in the vast majority of cases, and in the inflection of all new verbs entering the conjugation, it is retained in the perfect and third stem forms (*amā-re, amāv-ī, amāt-um*).

Table 3.6 Present, perfect and third stem forms of verbs of the *domō* type

<i>domāre</i>	<i>domuī</i>	<i>domitum</i>	‘tame’
<i>vetāre</i>	<i>vetuī</i>	<i>vetitum</i>	‘forbid’
<i>secāre</i>	<i>secuī</i>	<i>sectum</i>	‘cut’
<i>sonāre</i>	<i>sonuī</i>	<i>sonitum</i> (but fut. part. also <i>sonātūrus</i>)	‘sound’
<i>tonāre</i>	<i>tonuī</i>	<i>tonitum</i> (also <i>tonātum</i>)	‘thunder’
<i>crepāre</i>	<i>crepuī</i> (also <i>-crepāvī</i> in compounds)	<i>crepitum</i> (also <i>-crepātum</i> in compounds)	‘creak’
<i>cubāre</i>	<i>cubuī</i> (rarely <i>cubāvī</i>)	<i>cubitum</i>	‘lie down’
<i>lavāre</i>	<i>lāvī</i>	<i>lōtum</i> (also <i>lavātum</i>)	‘wash’
<i>iuvāre</i>	<i>iūvī</i>	<i>iūtum</i> (but fut. part. also <i>iuvātūrus</i>)	‘help’
<i>micāre</i>	<i>micuī</i>	[no form]	‘glitter’
<i>fricāre</i>	<i>fricuī</i>	<i>frictum</i> (also <i>fricātum</i>)	‘rub’
<i>plicāre</i>	<i>plicuī</i> (also <i>plicāvī</i>)	<i>plicitum</i> (also <i>plicātum</i>)	‘fold’
<i>necāre</i>	<i>necāvī</i> (but <i>-necuī</i> in compounds)	<i>necātum</i> (but <i>-nectum</i> in compounds)	‘kill’

(after Allen and Greenough 1903: §209)

Verbs of the *domō* type go against this very strong generalization. However, they do not do so by featuring inherently ‘irregular’ morphology, otherwise unknown to the verb system; rather, in the other parts of the paradigm they show inflectional behaviour which would be unremarkable for verbs belonging to conjugations other

than the first. The perfect formation which makes use of *-u-* built on to the lexical root (shown by *domuī*), with no stem vowel in evidence, is one of several available to third and fourth conjugation verbs, as well as being the major formation used by verbs of the second conjugation (*moneō monuī* ‘warn’); the formation seen in *lāvī*, *iūvī*, showing a lengthened vowel in the root, is likewise unknown to the first conjugation outside these forms, but not exceptional in terms of the verb system as a whole, cf. second conjugation *foveō fōvī* ‘cherish’, third conjugation *emō emī* ‘buy’. Meanwhile, the use of *-it-* or *-t-* (as in *domitum*, *sectum*) as the third stem marker is commonplace among non-first conjugation verbs.

Table 3.7 Inflected forms of heteroclite *domō* ‘tame’, in comparison with corresponding first and non-first conjugation forms

1 st conjugation ‘love’	(heteroclite) ‘tame’	2 nd conjugation ‘warn’
<i>amō</i>	<i>domō</i>	<i>moneō</i>
<i>amāre</i>	<i>domāre</i>	<i>monēre</i>
<i>amāvī</i>	<i>domuī</i>	<i>monuī</i>
<i>amātum</i>	<i>domitum</i>	<i>monitum</i>

The heteroclite nature of the paradigm of *domō* can thus be illustrated by juxtaposing it with both first conjugation and non-first conjugation forms, as in Table 3.7: the split between the *infectum* and the rest of the verb paradigm is evident. Given what has just been said, however, it is important to note that although I use forms of the second conjugation verb *moneō* as representative of distinctive non-first conjugation morphology here, it would be unwarranted to treat *domuī* and *domitum* as patterning with the second conjugation in particular, in the same way that *domāre* patterns unmistakably with the first conjugation; the crucial point is simply that they cannot be taken as first conjugation forms. The behaviour of the *domō* type thus differs from the instances of heteroclisism which have been touched on in this thesis so far, in that part of the morphology involved is best characterized *negatively*, as failing to belong

to a given class. Nonetheless, it seems uncontroversial that the label of heteroclisis does apply here: this case shows how the morphology in different parts of a single lexeme's paradigm can be recognized as incompatible, in the light of the system as a whole, even if it cannot all be assigned neatly to one existing inflectional class or another. This is one way in which heteroclisis can be less than fully canonical in the terms of section 2.3.

Many verbs of the *domō* type also show non-canonical behaviour of another kind, as can be seen in the full list given in Table 3.6, in that they do not consistently diverge from regular first conjugation behaviour in the perfect and third stems. The verb *plicō*, more usually found in compounds such as *explicō* 'unfold', *implicō* 'involve', optionally inflects without showing heteroclisis at all (*-plicāvī*, *-plicātum*), and *crepō* can also show regular first conjugation behaviour when part of a compound; meanwhile, *necō* as a simplex verb is almost always regular, only showing the non-first conjugation forms listed here in compound forms such as *ēnectus* 'exhausted'. And although, overall, an inflectional division within the paradigm can be identified as running between the first conjugation *infectum* and the non-first conjugation remainder, this division is sharper for some verbs than others. Thus *cubō* occasionally shows regular first conjugation perfect *cubāvī* (but no associated third stem **cubātum*); on the other hand, *tonō*, *fricō* and *lavō* optionally show third stem forms patterning with the first conjugation, while future participles in *-ātūrus* are found alongside non-first conjugation supines for *sonō* and *iuvo*.

No such variability appears in the inflection of the *infectum* forms of any of these verbs in Classical Latin, with one marginal exception: for *lavō* the language of poetry shows third conjugation as well as first conjugation forms in the *infectum*,

such as infinitive *lavere* alongside *lavāre*.⁷⁷ This is the retention of an archaic linguistic feature, as *lavō* with third conjugation morphology is found much more commonly, and in more naturalistic contexts, in pre-classical texts; and as documented by Bock (2008: 278-80), in Old Latin *lavere* and *lavāre* (and associated *inflectum* forms) are not merely morphological variants of a single lexeme, but represent different lexemes with different semantics, *lavere* meaning ‘wash’ and *lavāre* meaning ‘wash thoroughly, bathe’, a distinction effaced in the classical language. Meanwhile, pre-classical Latin also possessed third conjugation *sonere* (and associated *inflectum* forms) alongside *sonāre*, the latter eventually displacing the former only in the first century BC; individual attestations also exist (though they vary in reliability) of third conjugation infinitive *crepere*, present forms *percrepiti* and *tonimus*, and imperative *iuve*.⁷⁸

However representative each of these attestations may be of the genuine archaic situation, for Classical Latin the generalization can be made that verbs of the *domō* type do show regular first conjugation morphology in the *inflectum* – which they mostly fail to replicate in the rest of the paradigm. The question arises, then, to what extent the existence of parallel first and non-first conjugation *inflectum* forms in the earlier development of at least some *domō*-type verbs is connected with the group’s idiosyncratic behaviour in the classical language. I return to this issue below. However, I begin my treatment of the origins of the *domō* type by considering the way in which the Latin verb system came to possess the various conjugational characteristics found in the tables above, whose particular combination in the paradigm of *domō*-type verbs is synchronically so unexpected: this follows from the preliminary presentation given above in section 3.1. As will be seen, the emergence

⁷⁷ E.g. *dulcī mala vīnō lavere* ‘to wash evils in sweet wine’, Horace *Carmina* 3.4.61.

⁷⁸ For all these forms see the relevant sections of Bock 2008 as well as the discussion below.

of this form of heteroclisis can only be understood in the context of the structural and phonological development of the Latin verb in general.

3.3.1 Further background on the Latin verb system

Recall the Indo-European background to the tripartite Latin verb paradigm (cf. section 3.1.2), with its present, perfect and third stems. The *inflectum* forms of the verb ultimately continue the Indo-European ‘present’ or durative stem: this could be *basic* (‘primary’), i.e. consisting of the root itself or making use of one of several present-stem affixes, or *derived*, by means of a suffix which made a semantic or structural contribution, e.g. causatives, iteratives and denominals. The Latin *perfectum*, meanwhile, continues two distinct PIE stem categories, the (punctual) aorist and the (originally stative) perfect, which fell together as a perfective past tense around the time of the breakup of Italic into the Sabellic and Latino-Faliscan groups; where both aorist and perfect stems existed to a given root, each branch of Italic settled on one or the other as its perfect stem. The perfect was usually marked in the proto-language by initial reduplication; the aorist could consist of the root only or show a more complex structure, an aorist formation using the suffix *-s- apparently becoming favoured at a late stage. Finally, the third stem of Latin continues the verbal adjective in *-to- (> perfect passive participle) and the verbal noun in *-tu- (> supine), which were built directly on to the root.

The origin of Latin verb morphology in these PIE formations, which between them divide up the paradigm into three formal zones, is illustrated by the third conjugation verb *vehō -ere, vēxī, vectum* ‘carry’ from the PIE root *weg^h- (cf. *LIV*²: 661), which showed a thematic present (*wég^h-e/o-) and an s-aorist (*wég^h-s- / *wég^h-s-). Similarly, an inherited n-infix present and reduplicated perfect can be recognized in third conjugation *pangō -ere, pepigī, pāctum* ‘fix’, representing the

root **peh₂g-*. This is not to say that the original formations have remained morphologically untouched: to take just one example (cf. discussion of *finġō* ‘fashion’ on p. 147 above), the *n*-infix present, which was an athematic formation in the proto-language (thus 3sg **ph₂-ne-g-ti*, 3pl **ph₂-n-g-enti*), was recast as thematic (**pang-e/o-*) at the proto-Italic stage, with the result that *pangō* shares a conjugation with originally thematic stems such as *vehō*. However, even when remodelling of this kind is taken into account, it is easy to see how the variety of stem formations inherited from PIE, coupled with the retention of both aorist and perfect morphology to represent the *perfectum*, led to unpredictability in the relationship between the three Latin stems.

A different type of complication in the relationship between these stems can be attributed to a structural feature of the PIE system. Unlike in the Latin system which emerged from it, where relatively few verbs are defective, in PIE the existence of one aspect stem did not entail the existence of the others (Meiser 1998: 181; 2003: 76). In particular, as touched on above, it seems likely that derived stems existed only as part of the *present* system: no aorist or perfect counterparts existed specifically to these derived stems (as opposed to the primary present stem on which they were based). Similarly, there were no verbal nouns or adjectives corresponding to derived stems in particular, since these could only be formed directly to the verb root, as mentioned in section 3.1.2 above. Remnants of this principle can be detected in the Latin verb system, in cases where two separate verbs, differing in the *inflectum*, nonetheless share a single perfect and a single third stem, as in the case of *pendō* ‘weigh’ and *pendeō* ‘hang’ (both with perfect *pependī* and supine *pēnsum*) noted above; the same applies in the case of originally derived *sedeō -ēre* ‘be seated’ alongside basic *sīdō -ere* ‘sit down, settle (mostly of things)’ (< reduplicated present

**si-sd-*) which likewise agree in the perfect and supine (*sēdī*, *sessum*). With reference to the behaviour of the Latin perfect in particular, Meiser (1998: 179f.) labels this the *Simultanperfekt* phenomenon, a label I will adopt for convenience here.

However, at the same time, this feature of the PIE system is also responsible for the overwhelming *regularity* of the relationship between the stems of the Latin first conjugation, against which the behaviour of the verbs of the *domō* type stands out. In the Latin of all periods, as in the Sabellic languages (Buck 1904: 162), the large majority of first conjugation verbs stand synchronically in a derivational relationship with nouns (*cūrō* ‘care for’ cf. *cūra* ‘care’) or adjectives (*sānō* ‘heal’ cf. *sānus* ‘healthy’),⁷⁹ and these derived categories clearly made the principal lexical contribution in early Italic to what would become the first conjugation. Originally the two differed in form and showed only a coincidental resemblance (Meiser 1998: 186), as the denominal formation was originally thematic (*-*ā-ye/o-* <*-*eh₂-ye/o-*, with *-*eh₂-* originally corresponding to the nominal stem) while the deadjectival formation was athematic (*-*ā-* <*-*e-h₂-*), but these differences were eventually eradicated thanks to a combination of the general preference for thematic formations in Latin and the loss of intervocalic *y*. However, even before this point they shared the characteristic that, as derived formations, they had inherited no further stem forms. When this situation changed in proto-Italic and the paradigm of these verbs was filled out, this was done by abstracting a stem in *-*ā-* (i.e. the existing present stem without its distinctive present marker *-*ye/o-* where this occurred) and treating it as the basis upon which to build the rest of the paradigm. This ‘decharacterized stem’ in essence served as a newly created root upon which to build the stems used within the paradigm. The results of this process can be seen in both Sabellic (e.g.

⁷⁹ Alongside denominal and deadjectival derivatives, first conjugation verbs are also found which derive from existing verbs belonging to other conjugations: e.g. *occupō* ‘occupy’ from *capiō* ‘take’.

Oscan **fama-tted** ‘said’, showing the Oscan perfect in *-tt-*) and Latin (e.g. *cūrā-vī*, *cūrā-tum*), cf. Meiser (2003: 135).

The original status of the majority of first conjugation verbs as derived present stems, to which further stems were then created according to a uniform principle in Italic, explains why the first conjugation generally avoids the types of unpredictability over the paradigm which characterize the Latin verb system more generally. However, these regular stems, once created, were not exempt from further developments within Latin. As Meiser (2003: ch. 6) has argued, on both internal and comparative grounds it is likely that when an aorist was first created for these stems in *-ā-* it made use of the formation in **-s-*, which retained its productivity in most IE branches. This is true for the counterparts of these formations elsewhere (e.g. Ancient Greek aorist *etímēsa* < *etímāsa* from present *timáo* ‘honour’, derived from *tímē* < *tímā* ‘honour’; Old Irish aorist *mórais* from *mórain* ‘enlarge’ cf. *mór* ‘large’); moreover, the *s*-aorist formation is widely continued in Latin, in both inherited formations (such as *vēxī* to *vehō* above) and forms newly created within Latin (*iūnxī* to *iungō* ‘join’, which in PIE had a root aorist, cf. *LIV*²: 316). However, it is a notable feature of the Latin verb system as we find it that reflexes of the *s*-aorist are *never* attested on vocalic stems. Instead, perfect stems making use of a formant *-v-* are exclusively found (*cūrāvī*), which must at some point have come to be taken as the only acceptable formant following a vowel, and thus replaced earlier **s*. The origins of the *v*-perfect are unclear,⁸⁰ but this historical process is confirmed by the fact that it is also found in place both of ancient *s*-aorists inherited from PIE, e.g. *nēvī* (to *neō* ‘spin’) replacing **snēh₁-s-* / **sneh₁-s-*, and of original root aorists, e.g. *quiēvī* (to *quiēscō* ‘be calm’) replacing **k^wyeh₁-* / **k^wih₁-*.

⁸⁰ For discussion, see e.g. Meiser (2003: ch. 14), Seldeslachts (2001: 16-19); Rix 1992. The *tt*-perfect of Oscan and *ns*-perfect of Umbrian are similarly poorly understood.

In light of the behaviour of the first conjugation, where the identity of the stem vowel *-ā-* throughout is readily seen, less transparent patterns of stem alternation in other groups of verbs find an explanation too. The pattern shown by *moneō*, *monuī*, *monitum* ‘warn’ is prominent among second conjugation verbs, and yet the stem vowel *-ē-* of the *inflectum* is not evidently represented in the perfect or third stems. As the present stem itself exemplifies the derived causative formation of IE (**mon-eye/o-*, to root **men-* ‘think’), it would not originally have been accompanied by any other stems. Nonetheless, Sihler (1995: 586), following Sommer (1914: 559, 601), suggests that underlying the Latin third stem *monit-* is an inherited subsidiary stem in **-i-* (thus **mon-i-t-*), which also appears to be associated with the reflexes of the causative in Vedic Sanskrit. If that is the case, *monuī* could reflect **mon-i-w-* with vowel weakening, cf. *trīduum* ‘three days’ < **trīdiwom*. But as discussed at length by Meiser (2003: 136f.), there is good reason to think that before vowel weakening the third stem of these verbs in fact contained *-e-*. The verb *mereō* ‘earn’ behaves identically to *moneō*, its Classical Latin third stem being *merit-*; however, the vocalism of *meretrīx* ‘prostitute’, derived from this third stem, points to an earlier *mer-e-t-*, which is in fact attested in inscriptions (in the form *meretōd* ‘deservedly’: Weiss 2009: 438).

What is required is an account showing how proto-Italic could have innovated the elements **mone-*, **mere-* upon which these third stems are apparently built. However, the example of the first conjugation shows that this innovation is not unexpected but instead follows naturally as part of a wider trend. In order to fill out the paradigm of those derived second conjugation verbs which had no inherited connection with non-present stems (i.e. unlike *pendeō*, *pependī* above), proto-Italic abstracted a decharacterized stem in **-e-* on the basis of the causative-intensive

*-eye/o-,⁸¹ treating this as a complex element containing the present marker *-ye/o-, itself a familiar inheritance from PIE. The process was thus the same as that seen for the denominal verbs in *-ā-ye/o-. Subsequent vowel weakening gave third stem *monit-* < *-et-, as we have seen. Meanwhile, within the history of Latin the expected aorist formant -s- was replaced as a result of its post-vocalic context; vowel weakening in the resulting **mone-w-* gave the attested perfect *monuī*.

The preceding paragraphs have gone some way towards motivating the existence elsewhere in Latin of the inflectional behaviour which comes together incongruously in the paradigm of *domō* type verbs, i.e. regular first conjugation morphology as well as the correspondence between perfects and supines in *-uī* and *-itum* respectively (within the second conjugation). Two further points remain to be clarified, however – one accounting for the fact that the *-uī*, *-itum* pattern can be found outside the second conjugation at all, and the other, conversely, showing how its former presence in some second conjugation forms has been disguised by subsequent sound change. The latter concerns verbs whose decharacterized stem would have ended in *-we-, leading to a perfect stem in *-wew-. Such a sequence does not survive anywhere among the perfect forms of second conjugation verbs, and instead we find what can be described synchronically as long-vowel perfects, such as *fāvī*, *mōvī* (for earlier **fawewī*, **mowewī*) from *faveō* ‘favour’, *moveō* ‘move’. The phonetic details of the resulting forms, and their relationship with corresponding *fautum mōtum* etc., have been explained in various ways, as analogy as well as sound change is likely to have played a role (see Seldeslachts 2001: 7f. for discussion). However, the fact that long-vowel perfects are found in all such cases, rather than thinkable perfect forms in *-vuī, has consequences for our view of the *domō* type:

⁸¹ Cf. Meiser (2003: 136). The fact that this process affected verbs such as *mereō*, which continues an IE ‘essive’ in *-eh₁-ye/o- rather than a causative-intensive (*LIV*²: 570), shows that the two groups had already become conflated (> ‘the second conjugation’) in proto-Italic.

specifically, it means that the long-vowel perfects *lāvī* and *iūvī* should not be taken to represent a different structure diachronically from the perfects in *-uī* shown by the other heteroclite verbs in question.

Evidently, however, it is of special importance to an account of the heteroclite behaviour treated here to note that, even if the verbs of the *domō* type are left aside, *-uī* and *-itum* are not exclusively associated with the second conjugation, meaning that they do not always continue a decharacterized stem in *-e*. This is partly due to productivity in the *-uī* formation (or its ancestors) in particular. Sometimes it is clear that an attested perfect in *-uī* has been created secondarily, and relatively late in the development of Latin, on the basis of an existing perfect stem: this is seen in the adoption of classical *nexuī* in the place of earlier *nexī*, which was the expected outcome of the *s*-aorist of third conjugation *nectō* ‘weave’. Similarly, *messuī* as the perfect of third conjugation *metō* ‘reap’ must be a morphologically reworked version of **messī* < **met-s-* (Sihler 1995: 586). An earlier spread of *-uī* beyond its original scope can be seen in several other third and fourth conjugation perfects, such as *seruī* to *serō* ‘entwine’: since the inherited *s*-aorist **sēr-s-* / *ser-s-* (*LIV*²: 434f.) would regularly give **serrī*, the attested form of the perfect shows that (the phonetic forebear of) *-uī* replaced the formation in *s* before the prehistoric development **rs* > *rr*.

But these innovations in the perfect were not paralleled in the third stem forms, which show no sign of *-itum* (*nexum* < **nect-to-*, *messum* < **met-to-*, *sertum* < **ser-to-*). The behaviour showed by e.g. *vomō -ere -uī -itum* ‘vomit’ is thus not explained by the spread of productive *-uī* outside its original domain. Instead, the ancestry of this pattern is illuminated by comparative evidence that the PIE roots involved ended in a laryngeal. As described in 3.1.3 above, the third conjugation

infectum of *vomō* is explained by the thematization of the root present **wemh₁-* (/ **wmh₁-*), in which phonetically expected but morphologically anomalous **wema-* was replaced by **weme/o-*, most likely in proto-Italic. At the same time, however, the medial short *a* resulting from the laryngeal would have survived in other stems built upon this root. In this light, *vomitum* can be seen to continue **wema-to-*; *vomuī*, meanwhile, reflects the replacement of an earlier *s*-aorist of the form **wema-s-* with innovative **wema-w-*, in accordance with the principle that the *s* formation in Latin is not available after a vowel. Thus the agreement between *vomuī*, *vomitum* (built on **wema-*) and *monuī*, *monitum* (built on **mone-*) is secondary: it results from the identical treatment of short *a* and *e* in these contexts in the process of vowel weakening (Meiser 1998: 204).

3.3.2 *The origins of the domō type*

The existence of Latin perfect and third stem forms in *-uī*, *-itum* respectively, and their association with third and fourth but especially with second conjugation present stems, can thus be understood as the result of only a limited number of morphological and phonological developments on the basis of the PIE state of affairs. The existence of a regular first conjugation with a consistent stem vowel *ā* seems to emerge even more straightforwardly, primarily thanks to the generalization of an inherited third stem formation in *t* and an innovative perfect formation in *v* to verbs which previously only formed a present stem. However, what remains to be accounted for is the juxtaposition of these two patterns of behaviour within the conjugation of verbs of the *domō* type. Which pattern is original to these verbs, or (as with the *capiō* type) is it anachronistic to assume that either existing pattern should be taken as the starting point? And should any significance be ascribed to the

attestation of third conjugation *infectum* forms, in parallel with the more common first conjugation forms, for at least some of the items involved?

De Vaan 2012 addresses these questions in the course of an article devoted to the makeup of the Latin first conjugation more generally, a topic on which he promotes an unorthodox position. As noted above, the first conjugation predominantly continues two productive derived formations of PIE, namely the athematic deadjectivals in **-eh₂-* and the thematic denominals in **-eh₂-ye/o-*; but it is generally thought (see e.g. Meiser 1998: 189) that productive *deverbal* types arose from these later by extension: thus the ‘frequentative verbs’ in *-tāre/-sāre*, such as *cantāre* ‘incant’ in relation to *canere* ‘sing’, began life as a subtype of the deadjectival verbs built on perfect participles (*cantus* ‘sung’), while the so-called ‘intensives’, such as *occupāre* ‘occupy’ in relation to *capere* ‘take’, were originally based on prefixed nouns (such as **ob-caps* ‘occupier’, cf. *auceps* ‘bird-catcher’ < **avi-caps*). By contrast, de Vaan prefers an earlier hypothesis according to which, along with the denominal and deadjectival suffixes, Latin inherited a dedicated deverbal suffix **-āye/o-*⁸² with atelic semantics, which remained productive; this would be represented in both the ‘frequentatives’ and the ‘intensives’ but also in many other verbs, including some not generally ascribed to any particular structural category within the first conjugation, such as *cōnārī* ‘try’ or *mānāre* ‘spread’.

It thus seems that for de Vaan most first conjugation verbs which are neither deadjectival nor denominal are likely to reflect this proposed deverbal **-āye/o-*, whether or not a corresponding basic verb is also found. But where the requisite basic verb can still be identified in Latin alongside its first conjugation counterpart,

⁸² De Vaan reconstructs this suffix in the form **-āje-*, rather than resolving the long vowel into the sequence **eh₂* as is standard in PIE reconstruction. No significance should be attributed to this fact: the suffix intended is identical to the thematic denominal suffix which I have presented in the form **-eh₂ye/o-*.

de Vaan's case is particularly easily made. On this view, morphological doublets such as *lavere* – *lavāre* and *sonere* – *sonāre* represent instances of the phenomenon in question: the first conjugation forms are relative innovations produced on the basis of the earlier third conjugation forms by the productive application of this inherited atelic suffix, potentially late in the history of Latin.⁸³

De Vaan's account of the morphology of the *domō* type relies on the existence of pairs such as these. On the assumption that first conjugation *sonāre* was a latecomer, the forms *sonuī sonitum* must originally have been associated with basic *sonere*, but they then came to cater to both present stems simultaneously: this automatically produced heteroclisis in the paradigm of *sonāre*, as became more obvious once third conjugation *sonere* itself fell into disuse while the other forms originally associated with it survived. De Vaan attributes a similar history to the other 'deverbal *ā*-presents' which he observes to show forms in *-uī*, *-(i)tus* (his own list contains all the verbs given in Table 3.6, with the exception of *plicō*): '[w]e are free to assume that some perfects in *-uī*, *-itus* were made to a proto-Italic present (*gemere* – *gemuī*, *occulere* – *occuluī*, etc.) and stayed in place while the present was joined or ousted by a competing formation in **-āje-*' (de Vaan 2012: 329).

According to this explanation, the origin of the heteroclite behaviour displayed by verbs of the *domō* type (unlike that of the *capiō* type) owes nothing to sound change, but involves a discrepancy between the two parts of the paradigm which has always been genuinely morphological in nature. At the same time, heteroclisis did not arise here because one part of the paradigm of these verbs simply adopted the morphology of an alien inflectional class in place of its previous forms, as was the case for the Greek nouns treated in Karatsareas 2011a. Rather, the

⁸³ Cf. de Vaan (2012: 316): 'For instance, *sonāre* occurs beside and as a replacement of *sonere*, without any visible semantic difference. Thus, *sonāre* may well be a recent formation.'

innovative first conjugation forms found in the *infectum* did not emerge as inflectional variants of earlier non-first conjugation forms, but represent a distinct lexeme created by derivation: heteroclisis arose because in all these cases the new lexeme shared its perfect and third stems with the basic verb. The plausibility of this view is supported by the fact that *lavāre* and *lavere* do continue to show different semantics well into the attested history of Latin, as well as by the observation, made above, that Latin does possess other pairs of verbs which are clearly lexically distinct but share perfect and third stem forms (the *Simultanperfekt* phenomenon treated above), a structural feature inherited from PIE. The tendency for only the derived heteroclite verb to survive must reflect a loss of semantic distinctiveness between the two verbs in each pair, along with the status of the first conjugation as the principal productive class in Latin.

However, de Vaan's treatment of the *domō* type leaves some crucial questions unanswered. For one thing, as the lines cited show, the fact that all the verbs involved have a perfect in *-uī* (aside from *lāvī*, *iūvī*, dealt with above) is simply taken for granted: the article does not discuss the provenance of these forms, or the absence of any sign of the rival formation in *-sī* found widely in non-first conjugation perfect stems otherwise. More strikingly, this explanation does not address the question why these verbs in particular show heteroclisis. De Vaan recognizes that the majority of the verbs he is considering do show regular first conjugation morphology throughout the paradigm, identical to that shown by denominal and deadjectival verbs in *ā*, and this is the case whether or not the purported basic verb continues to exist (e.g. *dicāre -āvī -ātum* 'assign, dedicate' to *dīcere dixī dictum* 'say'); if Latin really does possess a large morphological class of 'deverbal *ā*-presents' which

includes the frequentatives and intensives (and others) as well as the *domō* type, as de Vaan suggests, it is unclear what led to this difference in behaviour.

In light of these considerations, I prefer a different account of the behaviour of the *domō* type, which acknowledges a common feature shared by many of the heteroclite verbs concerned here. As is generally recognized (including by de Vaan himself), comparative evidence from other branches of IE allows us to reconstruct a final laryngeal for the seven lexical roots continued by the verbs *domō*, *sonō*, *tonō*, *iuvō*, *lavō*, *vetō* and *secō*; meanwhile, there is nothing to rule out a final laryngeal in the roots underlying *micō*, *cubō* and *crepō*, which are less well attested across the family.⁸⁴ To take the single example of the verb *tonō* ‘thunder’, whose PIE root is reconstructed as **(s)tenh₂-*, the presence of a root-final laryngeal is indicated by Vedic *stanáyati* (< PIE iterative **(s)tonh₂-eye/o-*), whose short vowel in the first syllable can only be explained if it was originally followed by two consonants.⁸⁵ But the presence of these laryngeals should be expected to have had consequences for the shape of the Italic reflexes of these verbs as well. In particular, while the PIE iterative/causatives in **-eye/o-* normally retain this vocalism into Italic and thus enter the second conjugation, this pattern would automatically be disturbed by the vowel-colouring properties of the laryngeals, mentioned in fn. 62 above. Thus Eichner (1974: 58) notes that, by regular sound change (**h₂e > a*), the direct Italic cognate of Vedic *stanáyati* would be **tonaye/o-*. A present stem with this shape would be formally distanced from the majority of the inherited iterative/causatives; what is more, it could reasonably be expected to fall together with the large group of stems in **-āye/o-* or **-ā-* which together produced the first conjugation, very likely for

⁸⁴ I address the remaining verbs *plicō*, *fricō*, *necō* below.

⁸⁵ This reflects ‘Brugmann’s Law’ (Kuryłowicz 1927, Collinge 1985: 13-21). To justify the reconstruction of a root-final laryngeal for all these verbs would evidently take us too far afield.

reasons of sound change alone.⁸⁶ The behaviour of *tonāre* in the *inflectum* can thus be motivated on the basis of an iterative **(s)tonh₂-eye/o-* which is already attested elsewhere in IE.

As is laid out most comprehensively by Rix 1999, this observation can be extended from *tonāre* to cover the other laryngeal-final roots treated here, and doing so enables us to account not only for their assignment to the first conjugation in the *inflectum* but also for their synchronically anomalous perfect and third stems, as well as for the existence (in some cases) of rival present stems which do not belong to the first conjugation. Consider the conjugational pairing *sonere* – *sonāre*, corresponding to shared *sonuī sonitum*. This continues a root reconstructed by *LIV*² (p. 611) as **swenh₂-*, which formed a root present reflected in Sanskrit *svani-ti*. As seen above for *vomō* < **wemh₁-*, this root present **swena-* < **swenh₂-* would undergo thematization in Italic to **swen-e/o-* and thus give attested third conjugation *sonere*;⁸⁷ meanwhile, the parallel existence of first conjugation *sonāre* can be explained as the result of **swonaye/o-* < iterative **swonh₂-eye/o-*, exactly as was seen in the case of *tonāre*. But as these two present stems both belonged to the same PIE root, they would only ever have been associated with a single aorist, which in this case was an *s*-aorist (**swēnh₂-s-/ swenh₂-s-*, continued in Vedic *svānīt*). The expected outcome of this *s*-aorist in Latin is precisely the attested form *sonuī* shared by *sonere* and *sonāre*: it shows the usual replacement of the *s* formant after a vowel, with *sonu-* < **swena-w-* standing in place of inherited **swena-s-* < **swenh₂-s-*.

⁸⁶ Any differences between the phonetic outcomes of **-aye/o-* and **-āye/o-*, in particular, must have been very slight.

⁸⁷ **swe-* > **swo-* > *so-* is a regular sound change in Latin (cf. *soror* ‘sister’ < **swesōr*, Weiss 2009: 139), meaning that in this case the *e-* and *o-* grades of the two stem formations fell together automatically; however, in other cases (e.g. *tonere* – *tonāre*, *crepere* – *crepāre*) analogical levelling must be held responsible.

Similarly, their shared third stem seen in *sonitum*, with its short medial vowel, can now be recognized as reflecting the earlier presence of a laryngeal.

A similar process would affect laryngeal-final roots which formed PIE root aorists, rather than *s*-aorists, alongside two distinct present stems which survived into Latin – because here too a perfect stem in **-a-w-* would result. It is likely that just such a situation underlies the behaviour of perfect *iūvī* (reflecting PIE root **h₁ewh₂-*). Corresponding to this in the *inflectum* we find only first conjugation forms (*iuvāre*), with the single exception of third conjugation imperative *iuve*, which is found in one citation from the archaic dramatist Accius (second century BC) and whose linguistic reality is thus often doubted. But as noted by Bock (2008: 268-70), the shape of the verb *iuvāre*, with its initial yod, cannot be explained solely on the basis of the iterative **h₁owh₂-eye/o-* (which would give **uvāre*): instead this iterative must have been remodelled on a reduplicated present stem **h₁i-h₁éwh₂- > *(y)ewa- > *yowa- > *yuwa-*, which would be thematized as **yuw-e/o-*: that is to say, a third conjugation counterpart to *iuvāre* did previously exist in Latin. The associated PIE root aorist standing alongside these two present stems, **h₁ewh₂-* (attested in Vedic *ávit* ‘helped’) > **ewa-*, if likewise remodelled to **yuwa-*, would surface in Latin as a *v*-perfect owing to its final vowel: the resulting **yuwa-w-* underlies the attested form *iūvī*, just as **fawe-w-* gave *fāvī* as noted above.

In these cases, then, the association of a single Latin perfect with two distinct present stems, entailing heteroclisis in the paradigm of *sonāre* and *iuvāre*, can be understood as an example of the *Simultanperfekt* phenomenon in Latin, as indeed it is in de Vaan’s approach. But for de Vaan there is no way to explain why *sonāre* and *iuvāre* in particular would have shown this behaviour, when the majority of ‘derived *ā*-presents’ do not. If, instead, the existence of *sonāre/iuvāre* alongside *sonere/iuvere*

represents a direct continuation of the PIE state of affairs rather than an innovation dating to within the history of Latin, this question is resolved.

Rix 1999 explains the heteroclite behaviour of all the laryngeal-final verbs of the *domō* type in this way,⁸⁸ whether or not two rival present stem forms are actually attested in Latin. Thus first conjugation *domō*, from root **demh₂-*, can be traced back to an iterative **domh₂-eye/o-* also continued in Vedic *damáyati*; but Rix supposes that at an earlier stage Latin also possessed a basic present **demne-* reflecting PIE *n-* infix **dem-n-h₂-* (cf. Ancient Greek *dámnēmi*), with which the ancestors of non-first conjugation *domū* and *domitum* were originally associated. In this case, the former ultimately reflecting a root aorist **demh₂-* remade as **dema-w-* with later levelling of the vocalism to fit *domō*, and the same levelling would have affected *domitum*. But the assumption that these rival present stems survived into Italic in every case is unnecessary. As was seen by Meiser (2003: 137), a decharacterized stem **doma-* could be abstracted from **domaye/o-* < **domh₂-eye/o-* (exactly as **mone-* was abstracted from **mon-eye/o-*) in order to provide the morphological basis for the rest of the paradigm; the resulting innovative **doma-w-*, **doma-t-* would give the attested forms *domū* *domitum*. Again, however, recognizing the relevance of the root-final laryngeal in these verbs allows us to explain in a principled way both the fact that only a small number of verbs show this behaviour, and the particular shape (usually *-ū* *-itum*) taken by the non-first conjugation forms involved.

The reconstruction of specific formations (such as root presents, root aorists, or potentially neither) to individual verb roots in PIE is subject to debate – as is the likelihood that a given root provided two separate present stems which survived into

⁸⁸ Rix assigns *crepō* and *cubō* to this category – the first on the grounds of the Vedic Sanskrit aorist *akrapīṣṭa* (in which the *i* is taken to reflect the laryngeal). However, the Vedic form may represent a different IE root (so *LIV*²: 370). *Micō* is later added to the list by Meiser (2003: 138).

Italic and the potential extra-Italic relationships of some of the Latin verbs involved. As a result, the history of particular Latin verb paradigms and the directionality of influence between their various inflectional forms thus continue to be subject to debate. Precisely because analogical levelling and sound change have generally effaced the evidence from vocalism in the root which, if left undisturbed, would allow us to identify a given form as inherited (such as **dema-w-*) or innovative (such as **doma-w-*), which of these possibilities is to be preferred is not evident in every case: thus the perfect *vetuī*⁸⁹ could represent a reworking of a root aorist **weth₂-* (Isebaert 1994[1995]: 43, fn. 6), or an innovative **wota-w-* based on the decharacterized root **wota-* underlying *vetāre* < **woth₂-eye/o-* (Seldeslachts 1997: 2.616). Nonetheless, it does seem clear that the existence of the *domō* type in Latin, as a type, is originally explained by the combination of phonological and morphological effects seen here: roots ending in laryngeals could give rise to present stems in **-aye/o-* which fell together with the stems in **-āye/o-* and **-ā-* generally underlying the first conjugation, but by the time this conflation took place, aorist and third stems in a short vowel (> *-uī -itum*) were already established for these verbs.

Not every verb belonging to the *domō* type as we find it, however, can be accounted for in terms of the effects of an inherited laryngeal. While in some cases the point is disputed (see e.g. de Vaan 2012: 321 on *cubō*), in others there is no doubt about this. The verb *necō* ‘kill’ reflects the well-attested IE root **nek-* ‘perish’, in which a final laryngeal is ruled out (*LIV*²: 451f.),⁹⁰ and more pertinently it is not inherited at all but derived from the noun *nex* ‘murder, death’, meaning that regular first conjugation behaviour would be expected. Similarly, *plicō* ‘fold’ continues the root **plek-* ‘weave’ in some guise: it is generally taken as back-formed from prefixed

⁸⁹ Early attestations of this verb show initial *vo-* throughout the paradigm, and this became *ve-* by regular sound change in the second century BC; cf. early *vortō* ‘turn’ (Sihler 1995: 44).

⁹⁰ Latin *noceō* ‘harm’ continues the causative **nok-eye/o-* from this root.

verbs which are themselves derived from nouns, such as *supplicō* from *supplex*, which will explain both the apparently weakened vowel of the root and the verb's first conjugation affiliation, but does not account for the heteroclisis it shows. *Fricō* 'rub' appears to be derived from an earlier adjective **fricus* (related to *friō* 'crumble', itself ultimately from **b^hriH-* 'cut'), which again allows for no historical influence from a laryngeal.

It is unlikely to be a coincidence that these verbs are very similar to each other in shape; what is more, they are also similar in shape to *secō* 'cut' and *micō* 'glitter', at least the first of which certainly belongs to the original core of the *domō* type thanks to a root-final laryngeal. This formal similarity thus seems to have promoted the spread of *domō*-type heteroclisis to these lexemes where it did not originally belong. Furthermore, semantics may also have played a role. A semantic connection can be seen between 'cutting' and the further destructive physical manipulations of 'folding' and 'rubbing' (cf. what has just been said on the origins of *friō*); at the same time, it is easy to understand the influence of *secō* 'cut' on *necō* 'kill'.

3.3.3 Conclusion

Like the pattern of heteroclisis within the *inflectum* seen in the previous section, the heteroclite inflectional behaviour seen across the paradigm in verbs of the *domō* type owes a great deal to the effects of regular sound change. What were at one stage distinctive present stem forms such as **domaye/o-*, reflecting the presence of a laryngeal at the end of the ancestral root, would have fallen together automatically (or almost automatically) with the outcomes of stems in **-āye/o-*, which together with stems in **-ā-* underlie the regular first conjugation; but they did not thereby take on first conjugation inflectional morphology in the perfect and third stem, and as a

result they became characterized in synchrony by heteroclisis between first and non-first conjugation forms.

At the same time, however, individual instances of *domō*-type heteroclisis may reflect a more morphologically complex historical situation, traces of which are still visible in the Latin of our attestations in certain cases. Some of the IE roots which have been discussed here are continued in parallel conjugational formations in the *inflectum*, which, however, share the rest of their synchronic paradigm: this is the case, for example, for first conjugation *sonō -āre* and third conjugation *sonō -ere* in early Latin. As we have seen, while this situation is unusual for Latin it can be identified as continuing the normal state of affairs for PIE, in which a single root could stand as the basis of more than one present stem. Heteroclisis in such cases, which becomes starker and more canonical with the abandonment of the non-first conjugation verb, thus survives as a trace of a morphological feature of the proto-language which has overwhelmingly been lost in the making of the Latin verb system.

Finally, it is worth noting that we have a certain amount of evidence to suggest that the heteroclitite behaviour seen here, while resulting from the operation of sound change on stems of particular shapes, has been able to spread somewhat to new lexemes, apparently along formal or semantic lines. In the next section, we will see an instance of heteroclisis in Latin which, while its ultimate origin is hard to pin down, owes practically its entire distribution over the lexicon to productivity of this kind.

3.4 The *petō* type

Like those belonging to the *domō* type, Latin heteroclite verbs of the *petō* type are marked out by the fact that the conjugational behaviour of their *inflectum* forms is incompatible with that shown in the rest of the paradigm. However, while in the case of the *domō* type what is involved is best described as the juxtaposition of first and ‘non-first’ conjugation behaviour, here we can positively identify the two particular conjugations implicated: they are the third (*petere* cf. *regere*) and the fourth (*petīvī* *petītum* cf. *audīvī* *audītum*).

As has been touched on above, while a regular morphological relationship can be identified between the present, perfect and third stems in verbs of the first conjugation (and it is against this background that the *domō* type can be identified as heteroclite), this is not true over the Latin lexicon as a whole. In particular, there is no predictability in the assignment of perfect and third stem formations to a given verb with regular third conjugation morphology in the *inflectum*, as among such verbs a wide variety of formal relationships can already be found between these three stems, even if the *petō* type is left out of consideration.

It may therefore appear unwarranted to identify heteroclitism in the verbs treated here, when it would be possible simply to say that *-īvī -ītum* stand among the many formants available to third conjugation verbs: after all, they do not compromise any single regular pattern found otherwise, and to that extent are not out of place in the paradigm in the way that the *-uī -itum* of the *domō* type are. However, to take this position would be to miss the obvious respect in which these forms in *-īvī -ītum* do stand out from the general mass of the inflectional morphology used by verbs with a third conjugation present stem. While fourth conjugation verbs are not overwhelmingly uniform in their stem formation as first conjugation verbs are, new

recruits to the class strongly favour perfect and third stems in *-īvī -ītum*, which are also well represented among those ancestral fourth conjugation verbs which continue derived (denominal/deadjectival) stems in **-iye/o-*; and these can be identified as containing the stem vowel *ī* which underlies the corresponding *infectum* forms, thus paralleling the firm formal relationship between these parts of the paradigm which is seen in the first conjugation.⁹¹ Synchronically, then, forms in *-īvī -ītum* show characteristic fourth conjugation morphology. What distinguishes verbs of the *petō* type as heteroclite is the fact that this affiliation with the fourth conjugation is not shared across the *infectum* as well.

These are the same conjugations whose inflectional morphology also contributes to the heteroclite behaviour of verbs of the *capiō* type. However, as heteroclisism in the *capiō* type is solely a feature of the *infectum*, the two phenomena are separate and logically independent, and in fact it is possible for a verb to belong to both heteroclite types at once. This is the case for *cupiō* ‘desire’ and *sapiō* ‘taste (of)’: the existence of third conjugation morphology in some *infectum* forms of these verbs is anomalous given the fourth conjugation forms which appear both elsewhere in the *infectum* and in the rest of the paradigm.

The verbs involved here are listed in Table 3.8. In the previous section we saw that the pattern of heteroclisism which characterizes the *domō* type, in which the *infectum* stands opposed to the rest of the paradigm, originally emerged thanks to the effects of sound change on a set of verbs which were linked primarily in terms of their phonological shape (as was also the case for the *capiō* type); however, the heteroclite pattern was later extended slightly beyond its original domain, in a way

⁹¹ In fact, it is likely that the fourth conjugation stem vowel appears as (long) *ī* in the perfect and third stems precisely because of the analogical influence of the behaviour of *ā* in the first conjugation (Meiser 2003: 135). Ordinarily, the ‘decharacterized stem’ upon which new perfect and third stems were built would be expected to contain (short) *i* from **-iye/o-*, just as *monūi monitum* reflect *e* from **-eye/o-*.

that suggests that some role may have been played by formal and/or semantic considerations.

Table 3.8 Present, perfect and third stem forms of verbs of the *petō* type

<i>petō</i>	<i>petīvī</i>	<i>petītum</i>	seek, aim at
<i>quaerō</i>	<i>quaesīvī</i>	<i>quaesītum</i>	seek, ask for
[<i>quaesō</i>	<i>quaesīvī</i>	<i>quaesītum</i>	beg]
<i>cupiō</i> (but <i>cupere</i>)	<i>cupīvī</i>	<i>cupītum</i>	desire
<i>arcessō</i>	<i>arcessīvī</i>	<i>arcessītum</i>	summon
<i>laccessō</i>	<i>laccessīvī</i>	<i>laccessītum</i>	provoke
<i>capessō</i>	<i>capessīvī</i> (also <i>capessī</i>)	[no form]	take hold of, engage in
<i>incessō</i>	<i>incessīvī</i> (also <i>incessī</i>)	[no form]	attack
<i>facessō</i>	<i>facessīvī</i> (also <i>facessī</i>)	<i>facessītum</i>	do eagerly, accomplish
<i>sapiō</i> (but <i>sapere</i>)	<i>sapīvī</i> (also <i>sapūī</i>)	[no form]	taste (of); be wise
<i>rudō</i>	<i>rudīvī</i>	(? <i>rudīt-</i>)	roar, bray
<i>terō</i>	<i>trīvī</i>	<i>trītum</i>	wear down

In the present case it is also clear that considerations of shape and meaning have both played a role in the lexical distribution of this heteroclite behaviour: a simple list of the verbs in question brings out the existence of a cluster of lexemes belonging to the semantic sphere of ‘wishing to obtain’, connected with a second cluster making use of a common element *-ess-*, which seems in most cases to be a deverbative affix with ‘intensive’ or ‘conative’ value (*laccessō* cf. *laciō* ‘entice’, *capessō* cf. *capiō* ‘take’, etc.). On the other hand, the origin of the behaviour shown by these verbs is far from clear. In what follows I attempt to lay out what can be known about the developments leading to the Latin state of affairs.

Two of the verbs in Table 3.8, which neither show *-ess-* nor seem semantically connected to the ‘core’ of this group, can probably be separated from the rest in terms of their historical morphology as well. The verb *terō* ‘wear down’, together with its prefixed forms which generally inflect in the same way, shows idiosyncratic behaviour that makes it only a marginal example of the type under discussion here: to the extent that it makes sense to abstract a root *tr-* from *trīvī* *trītum* (which might thus be expected to have *infectum* counterparts **triō* **trīre*, cf.

sciō scīre scīvī scītum ‘know’), this does not precisely match the *ter-* seen in the third conjugation forms, and thus the incongruity between these parts of the paradigm is not fully captured by the label of heteroclisia. In its inflectional behaviour *terō* is arguably more closely comparable to several other third conjugation verbs with *v*-perfects containing a long vowel, which generally reflect root aorists (e.g. *sēvī* < **seh*₁- in relation to *serō* ‘sow’ < reduplicated present **se-sh*₁-, *LIV*²: 517f.), than it is to the other verbs treated here. But *trīvī* is unlikely to represent a root aorist of *terō*, which itself is usually taken to continue a root present **terh*₁- / **trh*₁-. Instead it is possible (following *LIV*²: 632) that this paradigm is historically suppletive, reflecting a secondary association between **terh*₁- in the *infectum* and a somewhat similar root with the approximate shape **trey(H)g*- in the third stem, which was extended to the *perfectum* within the history of Latin. On the other hand, many accounts have been proposed which aim to explain the forms of *terō* on the basis of a single PIE root, but the morphological or phonological developments required by each of these are rather *ad hoc*, leaving it hard to judge their merits.⁹² In any case, it is clear that the mechanisms underlying the behaviour of *terō* are not responsible for the existence of the *petō* type in general; nor is there reason to think that *terō* had any influence on the recruitment of new lexemes to the group.

The case of *rudō* ‘roar, bray’ is less clear-cut. In the *infectum* this verb is found throughout the classical period and continues a root present **reudH*- / *rudH*- shared widely across IE; this may have formed an *s*-aorist, attested in Old Avestan (*LIV*²: 508). However, in Latin no perfect is attested other than *rudīvī*, but this itself appears only once, in a text by Apuleius (second century AD); strictly speaking, no supine is found at any period, but Apuleius also provides one attestation of the

⁹² Cf. Seldeslachts (1997: 3.553-73), who deals with several proposals. His own solution, which posits a root **terh*₁y-, is plausible but likewise difficult to evaluate, as in his view no other root ending in laryngeal + *y* survives into Latin.

corresponding verbal noun *rudītus* ‘braying’, which patterns with the perfect stem. Meanwhile, the other fourth conjugation forms in Table 3.8 are all attested over long periods (beginning in the first century BC at the latest) and are well established as commonly or exclusively used in the relevant cells. These lone attestations from Apuleius are thus not entirely satisfactory as evidence for heteroclite inflection in *rudō*, and have been called into question especially since the language of Apuleius is known to show certain other idiosyncrasies in its inflectional morphology, including the assignment of verbs to conjugations (cf. Bock 2008: 354, 370, who points to his *praesāgāre* for *praesāgīre* and third conjugation subjunctive *sorbāmus* ‘let us drink’ for more usual second conjugation *sorbeāmus*).⁹³ What is more, as no *infectum* forms of *rudō* are found in Apuleius, it is unknown whether this may instead reflect a wholesale move of the verb into the fourth conjugation in his Latin (i.e. **rudiō -īre*).

However, there is no evidence for any such development elsewhere in our Latin attestations or in Romance, where *rudō* is inherited as a third conjugation verb *ruzer* in Occitan (cf. *REW* ad loc.). It is thus possible that *rudō* was genuinely heteroclite like the other verbs presented here. Either way, it is surely relevant that one of the semantic spheres associated with the fourth conjugation was that of animal noises, e.g. *hinniō* ‘neigh’ (cf. Leumann 1977: 556), and verbs of that type inflected in *-īvī -ītum*: whatever the timing and extent of the shift of *rudō* into the fourth conjugation, it is likely that its meaning played an important role.

There is no danger that the anomalous behaviour identified for the other verbs in Table 3.8 might be simply an artefact of our attestations in the same way. In most cases, the forms in *-īvī* and *-ītum* given here exist as counterparts to third conjugation *infectum* forms across all authors, and from the pre-classical period onwards: except

⁹³ Cf. Koziol (1872: 267-309), cited by Bock 2008, on inflectional innovation and neologisms in Apuleius more generally.

where otherwise noted, they are the only forms found for the perfect and third stems of these verbs. The split between the two parts of the paradigm is thus a genuine synchronic feature of the *petō* type.

In the heteroclite types treated in the previous two sections, the relevant conjugational split in the paradigm could be seen to emerge as the (reasonably direct) result of sound change as it affected groups of verbs which historically shared some formal characteristic, namely an inherited **-Cye/o-* sequence in the *capiō* type and a root-final laryngeal in the *sonō* type. We have seen that, to the extent that it makes sense to identify heteroclisis in the paradigm of *terō*, attempts have likewise been made to explain the split on the basis of differential sound change. However, no such explanation can be generalized to account for the existence of the whole *type* in question here, which (unlike the previous types discussed) does not display a central group of lexemes sharing an inherited shape that could have given rise to the observed synchronic split between third and fourth conjugation morphology for phonological reasons. Instead, the diachronic process involved appears to have been more directly morphological, involving the secondary adoption of fourth conjugation morphology in a paradigm which did not previously contain it.

The spread of fourth conjugation morphology into part of the paradigm can be detected straightforwardly in the case of *quaerō* ‘seek, ask for’, as the diachronic change has left synchronic traces in the lexeme family of this verb. The etymology of *quaerō* is disputed (cf. Bock 2008: 336f.) but ultimately it is not crucial to the observations made here: with Szemerényi (1960: 232-38) and (more tentatively) *LIV*²: 260, I assume that the present stem continues the compound stem **ko-h₂eis-e/o-* from root **h₂eis-* ‘seek’. The perfect and third stem forms *quaesivī*, *quaesitum* found in our Latin attestations are paralleled in the corresponding forms of verbs

derived from *quaerō*, such as *requirō -sivī -sītum* ‘ask again, enquire after’, and further derivations from those, e.g. *requisitiō* ‘examination’. However, some nouns which evidently derive historically from the synchronically basic verb *quaerō* show no sign of the unexpected *ī* which appears in parts of its inflectional paradigm – for example *quaestus* ‘acquisition, profit’, *quaestiō* ‘investigation’, and *quaestor*, the title of a class of Roman public official.

These relic forms, whose direct morphological link with the verb paradigm has been lost, provide evidence that the attested third stem *quaesīt-* is a morphological innovation, replacing an earlier **quaest-* no longer used in the inflection of the verb. What is more, the formal relationship between *quaerō* and a third stem **quaest-* would be the same as that still found in the paradigm of third conjugation *gerō gessī gestum* ‘carry’ (< **ges-*, cf. Bock 2008: 260f.) and *urō ussī ustum* ‘burn’ (< **h₁ews-*, cf. *LIV*²: 245): in these the present stem shows the regular rhotacism undergone by an inherited single *s* in intervocalic position around the fifth/fourth century BC (Weiss 2009: 150-52), while the *s* survives intact in the third stem, ‘protected’ by the following *t*, as well as in the perfect stem (which continues an *s*-aorist). What this suggests is that the paradigm previously took the form *quaerō *quaessī *quaestum*, a pattern which was subsequently disturbed by the adoption of *-ivī -ītum* from the fourth conjugation. This reconstruction is supported by the fact that rhotacism is absent from the perfect *quaesivī*: this form must continue an earlier **quaessivī*, whose geminate *ss* was regularly simplified following a diphthong, but only after rhotacism had ceased to operate (this is a well-attested source of intervocalic single *s* in Latin, cf. *causa* ‘cause’ < *caussa*, Sihler 1995: 222). The absence of rhotacism in *quaesītum* demonstrates the effect of analogical pressure from the perfect stem.

In *quaesivī*, then, it seems that fourth conjugation *-ivī* has been affixed to the pre-existing perfect stem rather than building an innovative perfect on the basis of the stem *quaer-*, just as *nexuī* (from *nectō*) results from the secondary addition of *-uī* to earlier perfect *nex-*, as seen in section 3.3.1 above. In this respect, *quaerō* stands with *terō* as a special case among the verbs I have ascribed to the *petō* type: it is not canonically heteroclite along a dimension which is sensitive to the presence of additional irregularity unrelated to the opposition in inflectional class affiliation itself, as *quaesivī quaesitum* would not be the regular forms expected in the paradigm of a fourth conjugation **quaerīō*.

However, the situation is complicated somewhat by the fact that alongside *quaerō* Latin does possess a verb *quaesō* ‘beg’ (earlier *quaessō*), i.e. exactly the ‘expected’ present stem corresponding to *quaesivī quaesitum* given the existence of heteroclisis in the paradigm. Diachronically this represents the PIE desiderative formation in **-se/o-* applied to the root seen in *quaerō*, pointing to an original meaning *‘want to seek’, cf. the relationship between *videō* ‘see’ and *vīsō* ‘visit’ (< *‘want to see’). In the Latin of our attestations *quaesō* is in fact defective, mostly surviving in the 1sg and 1pl present indicative as a politeness particle (‘please’). However, it may be that *quaesivī quaesitum* (and earlier **quaessī *quaestum*) did double duty completing the paradigms of both *quaerō* and *quaesō* at an earlier stage, when the latter operated more freely as a lexical verb: this morphological relationship with *quaes(s)ō* could have helped to promote the survival of the geminate *s* in *quaes(s)ivī quaes(s)itum*, which now makes *quaerō* distinctive among the verbs of the *petō* type.

On the other hand, I disagree with Meiser (2003: 126, fn. 48), who seems to suggest that innovative *quaes(s)ivī* may have been adopted partly because it avoids

homonymy with some present tense forms of *quaesō* (e.g. 1pl *quaesimus*). Numerous Latin verbs show partial homonymy between present and perfect, as the result of various historical developments which are treated in detail by Meiser himself (his *Einfaches Perfekt*, Meiser 2003: ch. 13) and there is little evidence for avoidance strategies of this kind; notably, *vīsō* has perfect *vīsī*, which is not remade as **vīsivī*. There is no reason to suppose that the avoidance of homonymy was particularly desirable in the case of *quaes(s)o* **quaes(s)ī*. The motivating factor for the adoption of fourth conjugation morphology in *quaerō* (/ *quaesō*) is clear from what has been said above with regard to the *petō* type as a whole: *quaerō* belongs squarely to those verbs identified above as relating to the sense ‘wishing to obtain’, which – along with the group of verbs characterized by *-ess-* – make up the majority of the *petō* type. If the special cases discussed above are set aside, there remains only *sapiō* which does not belong to either of these two groups.⁹⁴

What remains to be examined is how the association was established which links the adoption of fourth conjugation morphology to this particular semantic sphere, on one hand, and to this particular lexical shape, on the other. In principle, these would seem likely to be two unrelated principles influencing the lexical distribution of *petō*-type heteroclisis. However, the two do not in fact concern totally distinct groups of lexemes. The verb *arcessō* ‘summon’ has the relevant meaning as well as the relevant shape, and in turn *laccessō* ‘provoke’ is not far removed from *arcessō* in its semantics, meaning that the two identifiable clusters of verbs belonging to the *petō* type overlap somewhat. It thus seems likely that the incursion

⁹⁴ The presence of *sapiō* among heteroclite verbs of the *petō* type is hard to account for on any grounds: note that a perfect *sapuī* is found for this verb, as might be expected for it (cf. *rapīō* – *rapuī* ‘snatch’), but this only appears later than *sapivī*. The Oscan *sipus* ‘knowingly’, a relic perfect active participle, suggests that *sapivī* might in turn have replaced an earlier **sēpī* (Untermann 2000: 677). The anomalous presence of this verb here might suggest that it is in fact the originator of the *petō* type; but explanations for why this behaviour might have arisen in the paradigm of *sapiō* meet with the same objections as for *cupiō* below.

of fourth conjugation morphology into the perfect and third stem forms of the paradigm began as a characteristic of one of these groups and was later extended into the other as a result of this overlap. The question, then, is where this behaviour began.

Latin possesses only a few verbs with an *inflectum* in third conjugation *-essō*, almost all of which can be identified as belonging to the *petō* type and are thus listed above: the exceptions, *petessō* ‘strive after’ and *incipessō* ‘begin’,⁹⁵ are simply unattested outside the *inflectum*, meaning that there is no positive evidence allowing us to assign them to the *petō* type. This strong association may suggest that these verbs apparently bearing an affix *-ess-* (whose morphological status, however, remains to be discussed) constitute the original territory of the heteroclite behaviour treated in the present section, from which it spread later to verbs of ‘wishing to obtain’ which did not share this formal characteristic. It is potentially appealing to treat the unusual inflectional behaviour of these verbs as conditioned in the first instance by another morphological feature, namely the presence of deverbative *-ess-*.

However, the appeal of this approach must be tempered by the fact that our understanding of the origins of this element leaves no way to motivate a link between it and fourth conjugation morphology. The standard historical account of deverbative *-ess-*, provided by Leumann (1977: 555, largely following Thomas 1935), treats it as created analogically on the basis of a particular type of relic future formation in *-ss-* still attested in early Latin and restricted to first and second conjugation verbs (giving forms such as *amāssō* ‘I will love’ and corresponding infinitives in *-āssere*), which continued the desiderative **-se/o-* of the proto-language; the formal relationship between *amō -āre* and *amāssō* is thus preserved in the relationship

⁹⁵ Cf. *petō* ‘seek’ and *incipiō* ‘begin’. These verbs are also found in the forms *petissō*, *incipissō*; likewise, the form *capissō* is also seen in place of *capessō*.

between *petō -ere* and *petessō*, while the functional distinction came to involve conativity rather than simple futurity. In fact, it is open to question whether the differences between the two phenomena allow the diachronic relationship between them to run in the direction suggested by Thomas (some objections are raised by de Vaan 2011: 30f.). But ultimately, deverbative *-ess-* does seem to represent a more recently productive variant of the desiderative **-se/o-* seen above. What is more, it is likely that the group of verbs in *-essō* given above contains at least one which continues this earlier **-se/o-* directly. The verb *incessō* ‘attack’, although it echoes the others in question here, does not contain an affix *-ess-* at all, as there is no putative **incō* from which it could be derived; instead, it is most plausibly seen as reflecting **in-cēd-se/o-*, a desiderative in **-se/o-* corresponding to *cēdō* ‘depart’ (Bock 2008: 199).⁹⁶ The crucial point is that outside these forms in *-ess-*, whether it appears to represent a morphological unit or not synchronically in each case, there is no particular link between the Latin reflexes of PIE desideratives and fourth conjugation morphology: no inherited desideratives have entered the fourth conjugation in the *inflectum*, and it was noted above that originally desiderative *vīsō* has perfect *vīsī* rather than **vīsīvī* (third stem *vīsum* is shared with *videō*).

This may suggest that the behaviour in question did not originate among the verbs in *-essō* after all, and only spread to them secondarily. On the other hand, it will be seen that we are hardly any more confident about how heteroclisis could have arisen originally among the verbs of ‘wishing to obtain’, so our ignorance should not rule out the idea that the verbs in *-essō* were after all the original locus of the

⁹⁶ *Arcessō* ‘summon’ may also constitute a stem in **-se/o-* directly rather than one in the more recently productive *-ess-*; it also exists in the form *accersō* (potentially from **ad-kers-se/o-*, linking it with *currō* ‘run’) and there is no agreement as to which of these is the earlier variant. Seven possible etymologies are considered in Bock (2008: 173-76). If it is possible to treat *arcessō* as based on *arceō* + *-ess-*, formally the most obvious solution, we must be dealing with a more conservative meaning of *arceō* (e.g. ‘hold’) than that generally known for it in Classical Latin (‘ward off’).

heteroclis̄is discussed here. It is perhaps more telling, however, that while *arcessō* and *laccessō* consistently show a perfect in *-ivī* (like *petō*, *quaerō* and *cupiō*), the same cannot be said for the other verbs in *-essō*, for which ‘simple perfects’ in *-essī* are attested alongside fourth conjugation forms. Thus, for example, in classical prose we find 3pl pf. ind. *incessiv-ere* (Pliny the Elder, *Naturalis Historia* 33.53) but also 3sg ppf. ind. *incess-erat* (Suetonius, *Nero* 35.3) representing the perfect stem of *incessō*. Furthermore, it seems unlikely that a small group of relatively infrequent (and semantically rather various) verbs in *-essō*, if they were the first to develop heteroclis̄is of this kind, would have been able to influence the behaviour of the much more common verbs of ‘wishing to obtain’, *petō*, *quaerō* and *cupiō*, which did not share the distinctive shape of the original group.

It has thus been widely accepted that the direction of influence was the reverse: the existence of heteroclis̄is between third and fourth conjugations in the verbs of ‘wishing to obtain’ led to heteroclis̄is of the same kind in the group of verbs characterized by *-essō*, whether or not this represented a discrete morphological element, by means of the lexical ‘bridge’ formed by *arcessō* and *laccessō*. Meanwhile, we have seen that *quaerō* also took on its (near-)heteroclitite behaviour secondarily. That leaves only *petō* and *cupiō* as potential starting points for this conjugational pattern. Almost all⁹⁷ consider that *cupiō* represents the more plausible alternative (endorsed by e.g. Sommer 1914: 565, Leumann 1977: 594, Schrijver 1991: 395, Seldeslachts 1997: 2.401), given the synchronic relationship and occasional fluctuations that exist in Latin between the *cupiō* type and the fourth

⁹⁷ Strangely, de Vaan (2011: 27f.) aims to cast doubt on the antiquity of perfect *cupivī*, noting that it is not attested until its appearance in Catullus (first century BC), ‘considerably later than the perfects in *-ivī* of *petere*, *quaerere* and *sapere*, for which *cupivī* is supposed to have provided the model.’ I fail to see the relevance of this, as there is no suggestion that at an earlier stage a different perfect formation is attested for *cupiō* instead. What is more, the perfect participle *cupītus*, itself displaying fourth conjugation morphology, is already attested in Plautus over a hundred years before, as de Vaan recognizes: as this represents the expected counterpart of a perfect *cupivī*, I see no benefit in supposing that the latter is a more recent formation.

conjugation. In particular, Meiser (2003: 237) suggests that *cupīvi* entered the paradigm of *cupiō*, (replacing an earlier **cucupī*, **cupsī* or the like) from the perfect **concupīvi* of a compound verb lost by the time of our attestations, which as a prefixed form belonged securely to the fourth conjugation as a result of the same differential treatment which produced fourth conjugation *amiciō* < **amb-jak-* (cf. fn. 74). However, the required compound verb **concupiō* is itself hypothetical; what is more, fourth conjugation present stems with this origin are not automatically accompanied by forms in *-īvi -ītum*, cf. *amiciō* itself (*amicuī*, *amictum*).

In fact, a very different solution has recently been proposed. Martzloff (2006: 317) notes, following Meillet 1922, that the paradigm of *poscō* ‘demand’ (< **prk-ské/ó-* from the root **prek-* ‘ask’) was defective in Latin with respect to its third stem: as a result, in the classical period, we find the relevant forms of *petō* ‘seek’ used instead. However, in Livius Andronicus (third century BC) we find an archaic supine *procitum* meaning ‘to woo’ and standing in relation to the noun *procus* ‘suitor’: the length of the *i* is uncertain, but if, with Watkins (1971: 60), this reflects a fourth conjugation denominal verb **prociō* (cf. *serviō* from *servus*), the form would in fact be *procītum*. However, this and *procus* both also reflect the root **prek-*, and are likely to have had more general meanings at an earlier stage. It is possible, then, that *procītum* was previously used to supplement the paradigm of *poscō* in just the same way as *petītum* is used in the Latin of our attestation. If this is the case, the unusual relationship between *procītum* and the rest of the paradigm of *poscō*, also a verb of ‘wishing to obtain’, could have been the starting point for the spread of fourth conjugation morphology into other verbs in this semantic field, perhaps *petō* first of all.

This proposal is evidently as speculative as any which envisages the heteroclite behaviour treated here as beginning in the paradigm of *cupiō*, but although it relies on a number of assumptions about features of the history of Latin we have no way of knowing, it is not inherently implausible; rather, it raises the general possibility that none of the solutions so far proposed are correct and indeed that the true starting point for the *petō* type may not survive in the Latin of our attestations at all. It is certainly probable that the true chain of events was just as intricate (and just as unlikely to convince, in the absence of appropriate evidence) as has yet been suggested.

Whatever the full history of this type of heteroclis, however, it clearly differs from the previous two cases treated in this chapter with regard to both its ultimate origin and the reasons for its lexical distribution. We saw above that heteroclis of the *domō* type can be identified as spreading to a certain extent, taking in certain lexemes which would not have been expected to display it on diachronic grounds, and tentatively identified a role for formal and semantic similarity in this spread; however, the majority of the verbs involved there acquired their heteroclite behaviour ‘naturally’, as a result of the contingent effects, specific to them, of phonological and morphological changes applying more generally in the development of the language. Similarly, the *capiō* type emerged as a distinct type as the result of sound change on the basis of an important existing formation. However, heteroclis of the *petō* type seems likely to have arisen at first in only a single paradigm, which may not in fact be any of those known to us to show the behaviour in question, and as the result of lexical contamination or morphological analogy rather than sound change; beyond that, its presence in Latin is due entirely to its

productive spread, which, however, largely involved only two well-defined groups of lexemes, one linked by semantics ('wishing to obtain') and the other by shape (the *-ess-* element).

4. Heteroclisis in Romanian

4.1 Introduction and overview of the Romanian verb

This chapter looks at the origins and development of heteroclisis as it has emerged, several times independently and for a number of different reasons, in the history of the Romanian dialects. Specifically, I deal here with varieties belonging to Daco-Romanian, one of the four dialect groups comprising the Romanian branch of Romance:⁹⁸ the varieties in this group, spoken in an area largely coinciding with present-day Romania and Moldova, include the standard Romanian language, whose verb inflection is the focus of much of this chapter. As will be seen, heteroclisis is by no means a marginal phenomenon in Romanian, but has figured a great deal in the evolution of the conjugational system, both in the dialects and in the standard language. I begin with an overview of Romanian verb inflection, before turning to the various types of heteroclite behaviour attested in the language; first among these I present the phenomena treated in detail in Maiden 2009, which emphasizes the importance of heteroclisis in the study of Romance morphology and thus represents a fitting starting point for the Romance portion of the present work.

While the verb systems of the Daco-Romanian varieties show differences on numerous points of detail, some of which will be noted in due course, for our purposes here the fundamental characteristics of synthetic verb inflection across the conjugations in this dialect group can be illustrated by representative paradigms from standard Romanian; the standard is also phonologically conservative in a crucial

⁹⁸ The other dialect groups making up the Romanian branch are Aromanian (Macedo-Romanian), spoken in communities scattered widely over northern Greece, Albania and the Republic of Macedonia; Megleno-Romanian, whose speakers are localized to a small region straddling the border between Greece and the Republic of Macedonia; and Istro-Romanian, spoken only in the Istrian peninsula of Croatia. In what follows, the term 'Romanian' can be taken to refer narrowly to Daco-Romanian unless otherwise specified.

respect which means that its behaviour can effectively be treated as ancestral to that seen in the varieties under discussion in section 4.2. In the tables which follow (based on Maiden 2009: 62f.), I give (the synthetic portion of) the inflectional paradigm of regular verbs belonging to the four conjugations generally recognized for Romanian, which correspond historically to the four conjugations of Latin.⁹⁹ These provide the context within which the instances of heteroclisis treated here are to be viewed.

A word should first be said about the status and composition of the different conjugations. The third and, especially, the second are small and almost entirely unproductive; the distinction I draw between two ‘types’ of the former (Tables 4.3 and 4.4) will be explained shortly. Both fourth and first conjugations are productive – the latter more so in the present day, while up to the 19th century the former predominated (Sánchez Miret 2006: 34-45; Jordan 1935: 50-64). The fourth conjugation has two subclasses, one of which bears a formative, or ‘inflectional augment’, *-esc-* (with allomorphs *-easc-*, *-eșt-*) between the root and the inflectional ending in certain cells; most fourth conjugation verbs make use of this augment. The first conjugation also has two subclasses, one of which (not illustrated here) makes use of an augment *-ez-* (*/-eaz-*) distributed identically to fourth conjugation *-esc-*: in the first conjugation, augmented and unaugmented verbs are much more evenly matched in number, but again the augmented type is the more productive (Maiden 2009: 62; Nedelcu 2013: 20f.). This phenomenon of inflectional ‘augmentation’, which is found widely across Romance and whose Romansh counterpart I examine in more detail with reference to the particular variety treated in section 5.4, is one of the most salient features of Romance verb morphology in contrast to that of Latin

⁹⁹ In due course (section 4.4) I address the question of a potential fifth conjugation which is also frequently identified in descriptions of the Romanian verb system; the distinctive morphological behaviour involved is intimately bound up with heteroclisis.

(Meul 2013: 1), and results from the generalization of a pattern which originally emerged from the falling together of pairs of derivationally related Latin lexemes to give a single inflectional paradigm with stem alternants drawn from both lexical sources. The fourth conjugation *-esc-* here thus reflects an element of the form *-ēsc-* as seen in *flōrēscō* alongside *flōriō*, both meaning ‘flower, flourish’ in Late Latin, while *-ez-* continues a borrowed Greek verbalizing suffix of the form *-iz-*, generally reconstructed for early Romance in the form **-idj-*. On all these points cf. Blaylock 1975, Maiden 2004, and especially Meul 2013.

Table 4.1 First conjugation (without augment), *a cânta* ‘sing’¹⁰⁰

Infinitive	<i>cântá</i>
Gerund	<i>cântánd</i>
Past participle	<i>cântát</i>

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>cânt</i>		<i>cânt</i>	<i>cântám</i>	<i>cântái</i>	<i>cântásem</i>
2sg	<i>cánți</i>	<i>cántă</i>	<i>cánți</i>	<i>cântái</i>	<i>cântáși</i>	<i>cântáseși</i>
3sg	<i>cántă</i>		<i>cánte</i>	<i>cântá</i>	<i>cántă</i>	<i>cântáse</i>
1pl	<i>cantám</i>		<i>cantám</i>	<i>cântám</i>	<i>cântărám</i>	<i>cântáserám</i>
2pl	<i>cantáți</i>	<i>cântáți</i>	<i>cantáți</i>	<i>cântáți</i>	<i>cântărăți</i>	<i>cântáserăți</i>
3pl	<i>cántă</i>		<i>cánte</i>	<i>cântáu</i>	<i>cântără</i>	<i>cântáseră</i>

Table 4.2 Second conjugation, *a putea* ‘be able’

Infinitive	<i>puteá</i>
Gerund	<i>putánd</i>
Past participle	<i>putút</i>

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>pot</i>		<i>pot</i>	<i>puteám</i>	<i>putúi</i>	<i>putúsem</i>
2sg	<i>póți</i>		<i>póți</i>	<i>puteái</i>	<i>putúși</i>	<i>putúseși</i>
3sg	<i>poáte</i>		<i>poátă</i>	<i>puteá</i>	<i>putú</i>	<i>putúse</i>
1pl	<i>putém</i>		<i>putém</i>	<i>puteám</i>	<i>putúrám</i>	<i>putúserám</i>
2pl	<i>putéți</i>		<i>putéți</i>	<i>puteáți</i>	<i>putúráți</i>	<i>putúserăți</i>
3pl	<i>pot</i>		<i>poátă</i>	<i>puteáu</i>	<i>putúrá</i>	<i>putúseră</i>

Table 4.3 Third conjugation (u-type), *a țese* ‘weave’

Infinitive	<i>țése</i>
Gerund	<i>țesánd</i>
Past participle	<i>țesút</i>

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>țes</i>		<i>țes</i>	<i>țeseám</i>	<i>țesúi</i>	<i>țesúsem</i>
2sg	<i>țeși</i>	<i>țese</i>	<i>țeși</i>	<i>țeseái</i>	<i>țesúși</i>	<i>țesúseși</i>
3sg	<i>țese</i>		<i>țeásă</i>	<i>țeseá</i>	<i>țesú</i>	<i>țesúse</i>
1pl	<i>țesem</i>		<i>țesem</i>	<i>țeseám</i>	<i>țesúrám</i>	<i>țesúserám</i>
2pl	<i>țeseși</i>	<i>țeseși</i>	<i>țeseși</i>	<i>țeseáți</i>	<i>țesúráți</i>	<i>țesúserăți</i>
3pl	<i>țes</i>		<i>țeásă</i>	<i>țeseáu</i>	<i>țesúrá</i>	<i>țesúseră</i>

¹⁰⁰ I cite Romanian verbs in the traditional citation form, namely *a* + infinitive, the equivalent of ‘to VERB’; for the English gloss I give just the bare verb form. The stress marks given on the verb forms presented in this chapter are not found in Romanian orthography. The letters *ț*, *ș*, *j* and *ă* represent [ts], [ʃ], [ʒ] and [ə] respectively, while both *â* and *î* represent [i]. The sequences *ea* and *oa* represent diphthongal [ɛa] and [ɔa]; unstressed *i* in final position is usually asyllabic.

Table 4.4 Third conjugation (s-type), *a arde* ‘burn’

Infinitive *arde*
 Gerund *arzând*
 Past participle *ars*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>ard</i>		<i>ard</i>	<i>ardeám</i>	<i>arséi</i>	<i>arsésem</i>
2sg	<i>ârzi</i>	<i>ârzi/arde</i>	<i>ârzi</i>	<i>ardeái</i>	<i>arséși</i>	<i>arséseși</i>
3sg	<i>arde</i>		<i>árdă</i>	<i>ardeá</i>	<i>ârse</i>	<i>arsése</i>
1pl	<i>árdem</i>		<i>árdem</i>	<i>ardeám</i>	<i>ârserăm</i>	<i>arséserăm</i>
2pl	<i>ardeți</i>	<i>ardeți</i>	<i>ardeți</i>	<i>ardeáți</i>	<i>ârserăți</i>	<i>arséserăți</i>
3pl	<i>ard</i>		<i>árdă</i>	<i>ardeáu</i>	<i>ârseră</i>	<i>arséseră</i>

Table 4.5 Fourth conjugation (without augment), *a dormi* ‘sleep’

Infinitive *dormi*
 Gerund *dormind*
 Past participle *dormit*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>dorm</i>		<i>dorm</i>	<i>dormeám</i>	<i>dormii</i>	<i>dormisem</i>
2sg	<i>dórmi</i>	<i>dórmi</i>	<i>dórmi</i>	<i>dormeái</i>	<i>dormiși</i>	<i>dormiseși</i>
3sg	<i>doárme</i>		<i>doármă</i>	<i>dormeá</i>	<i>dormi</i>	<i>dormise</i>
1pl	<i>dormim</i>		<i>dormim</i>	<i>dormeám</i>	<i>dormirăm</i>	<i>dormiserăm</i>
2pl	<i>dormiți</i>	<i>dormiți</i>	<i>dormiți</i>	<i>dormeáți</i>	<i>dormirăți</i>	<i>dormiserăți</i>
3pl	<i>dorm</i>		<i>doármă</i>	<i>dormeáu</i>	<i>dormiră</i>	<i>dormiseră</i>

Table 4.6 Fourth conjugation (with augment), *a citi* ‘read’

Infinitive *citi*
 Gerund *citind*
 Past participle *citit*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>citesc</i>		<i>citesc</i>	<i>citeám</i>	<i>citii</i>	<i>citisem</i>
2sg	<i>citești</i>	<i>citește</i>	<i>citești</i>	<i>citeái</i>	<i>citiiși</i>	<i>citiseși</i>
3sg	<i>citește</i>		<i>citească</i>	<i>citeá</i>	<i>citii</i>	<i>citise</i>
1pl	<i>citim</i>		<i>citim</i>	<i>citeám</i>	<i>citirăm</i>	<i>citiserăm</i>
2pl	<i>citiiți</i>	<i>citiiți</i>	<i>citiiți</i>	<i>citeáți</i>	<i>citirăți</i>	<i>citiserăți</i>
3pl	<i>citesc</i>		<i>citească</i>	<i>citeáu</i>	<i>citiră</i>	<i>citiseră</i>

While the synthetic paradigm in Romanian is notably smaller than that found in Latin, Romanian does preserve much of the inflectional morphology of the Latin active voice, and overall the distinctions between the different conjugations can be traced back to features of the ancestral system; at the same time, many formal and functional innovations have evidently taken place. In terms of the structure of the

paradigm it is worth noting that little remains of the systematic Latin opposition of indicative and subjunctive, orthogonal to tense: the Romanian pluperfect is descended from the Latin pluperfect subjunctive (e.g. *cântásem* < *cantāvissem*), while the present subjunctive category of Latin is continued by forms which are functionally tenseless (they are used in subordination regardless of the tense of the grammatical context: Maiden 2011a: 62), and accordingly will be referred to simply as the ‘subjunctive’ here. In fact, however, the first and second person present subjunctive forms of Latin have nonetheless been replaced by the corresponding indicative forms (except in *a fi* ‘be’), meaning that only the third person subjunctive forms remain morphologically distinct from the present indicative in Romanian (Zafiu 2013: 28f.); these always show syncretism between singular and plural, as would be expected on historical grounds (e.g. first conjugation *cánte* < *cantet*, 3pl *cánte* < *cantent*; non-first conjugation 3sg *țeásă* < *texat*, 3pl *țeásă* < *texant*).

Third person syncretism is likewise found in the first conjugation present, again reflecting the Latin situation: 3sg *cántă* < *cantat*, 3pl *cántă* < *cantant*.¹⁰¹ This is a distinctive characteristic of the first conjugation: elsewhere the 3sg present is characterized by *-e* and the 3pl usually by the absence of any inflectional ending, as a result of the generalization of originally third conjugation *-unt* to all non-first conjugation verbs (3pl *pot*, *dorm* < *potu*, *dormu* < **potunt*, **dormunt*), cf. Zamfir (2005: 429). In the first conjugation the 2sg imperative is also syncretic with the third person present forms in *-ă*. Outside the first conjugation, meanwhile, thanks to a striking redistribution of the non-first conjugation imperative morphology inherited

¹⁰¹ Feldstein 2004 distinguishes the third person syncretism seen in the subjunctive, which he considers a genuinely morphological feature, from that seen in the first conjugation present indicative, which he treats as a phonological effect resulting from the deletion of an underlying final number marker *-u* in the plural form. However, there is no evidence for an underlying *u* in 3pl *cántă*, nor does it reflect a segment which has ever existed distinguishing the two third person forms in the first conjugation: the origins of syncretism in subjunctive *țeásă* and indicative *cántă* are identical.

from Latin,¹⁰² the form taken by the 2sg imperative is associated with transitivity: as a general rule, for transitive verbs this is identical to the 3sg present, thus *țese* ‘weave!’ and *ârde* ‘burn! (tr.)’, while in intransitives it is syncretic with the 2sg present (/subjunctive), thus *dórmí* ‘sleep!’ and *ârzi* ‘burn! (intr.)’. The last of these examples also illustrates the innovative presence of final *-i* across all conjugations in the 2sg present/subjunctive (/imperative); in this context many consonants undergo palatalization or assibilation, in this case of *d* to *z* (cf. *cânt-í* from *cânt-á*).

Across much of the rest of the paradigm, conjugational distinctions are marked by means of (reasonably) consistent theme vowels, whose origins in the Latin system are clear. The theme vowel of the first conjugation is *a*, which is particular to this conjugation and appears, always under stress, in the infinitive, past participle, imperfect, preterite, pluperfect, and the 1pl¹⁰³ and 2pl present/subjunctive forms (the 2pl imperative is also syncretic with the latter in all verbs); however, the vowel *â* seen in the first conjugation gerund is held in common with the second and third conjugations. The behaviour of the fourth conjugation in these cells is very similar: this bears a distinctive stressed theme vowel *i* in all the parts of the paradigm just mentioned, including the gerund – with the exception of the imperfect, where conversely it is the fourth conjugation that patterns with the second and third and against the first, having borrowed *-ea-* < *-ēbā-* in place of inherited *-ia-* (still attested in the 16th century, cf. Rosetti 1968: 546) < *-ībā-* (which replaced *-iēbā-*).

The inherited distinction between the second and third conjugations is preserved more successfully in Romanian than in many other branches of Romance,

¹⁰² Nielsen Whitehead (2012: 226-32), who notes this association with transitivity on the basis of the figures in Diaconescu 1999, suggests that this pattern was promoted by the continuation of Latin final *-ē* as Romanian *-i* (rather than *-e* as has previously been assumed) and the tendency for second conjugation verbs to be intransitive.

¹⁰³ In the 1pl present/subjunctive it appears in the form *ă*. It is also possible to treat the unstressed *-ă* of the 3sg/3pl present (/2sg imperative) as a manifestation of this theme vowel, and to treat the corresponding 3sg present (/2sg imperative) *-e* of the other conjugations likewise (Iliescu and Mourin 1991: 138-40).

where the two are entirely or almost entirely conflated, chiefly owing to similarities in their vocalism and their common status as unproductive classes. Nonetheless, in Romanian too they pattern much more closely with each other than with the first or the fourth.¹⁰⁴ In the infinitive, the 1pl present (/subjunctive) and the 2pl present (/subjunctive/imperative), both conjugations can be taken to make use of a theme vowel *e* which, however, only bears stress in second conjugation forms (infinitive in *-eá*); the corresponding third conjugation forms are rhizotonic, reflecting the difference in vowel length (and thus stress placement) in the relevant Latin forms (*putém* < **potēmus* vs. *țésem* < *teximus*). The two also behave alike in showing a formal divide in the paradigm which is not seen in the first and fourth conjugations. In the preterite, pluperfect and past participle, the second conjugation and the majority of third conjugation verbs (labelled the ‘u-type’ above) take a stressed theme vowel *u* rather than *e*; the remaining third conjugation verbs (‘s-type’) show allomorphy in the root, which usually employs *s* in the preterite and pluperfect and *s* or *t* in the past participle. The relatively few ‘s-type’ verbs with past participles in *t* (e.g. *a rupe* ‘tear’, *rupséi*, *rupt*) constitute the only exception to the generalization that verbs make use of the same stem in the past participle as in the preterite and pluperfect (Zafiu 2013: 33).

These two patterns shown by second and third conjugation verbs in the preterite, pluperfect and past participle evidently reflect Latin *perfectum* forms in *-uī*, together with analogically created third stem *-ūtus* (Laurent 1999: 92-94), and in *-sī*, together with associated *-sus/-tus*. The formal split in the paradigm thus embodies the widespread morphomic pattern shown by Romance verbs, labelled the ‘PYTA’

¹⁰⁴ Moreover, while the conjugations remain distinct as morphological templates in Romanian, individual verbs assigned to them show a tendency to equivocate between the two types of inflectional behaviour, particularly in favour of the third conjugation forms (Nedelcu 2013: 20). There has also historically been a great deal of movement towards the third conjugation; note that *a arde* itself continues the Latin second conjugation verb *ārdeō -ēre*.

pattern (Maiden 2001) and introduced in section 2.5 above, whereby cells continuing the Latin *perfectum* share a distinctive stem marking them out from the continuants of the *imperfectum* which make up the majority of the paradigm; here, as often in Romance, the past participle aligns itself with the PYTA forms.

Another morphological split in the paradigm attested widely across Romance is also visible in the paradigms provided. In the ‘N-pattern’ (also introduced in section 2.5) as it affects the first, second and fourth conjugations, a stem may be found in the singular and third plural forms of the present indicative and subjunctive which is different from that seen elsewhere: this ultimately reflects the fact that, in these conjugations in Latin, the forms in these cells already differed from the remainder in that they bore rhizotonic stress, which often led to distinctive vowel quality in the Romance descendants of the relevant forms. This pattern can be seen in the alternation of *o* and *u* in the paradigm of *a putea* (modified in some forms by the further development of *o* to *oa*: see immediately below) and in the distribution of the augments *-esc-* and *-ez-*. (In the third conjugation, where a distinctive N-pattern stem exists it is shared by the infinitive and all present and subjunctive forms, all of these being rhizotonic in Romanian as in Latin.)

The survival in Romanian of a further morphological pattern affecting stem shape (the ‘L-pattern’) which is inherited in common by most Romance languages, but only obliquely represented in the tables above, will be dealt with in due course (section 4.3). I close this section, however, by pointing out the effect on some verb paradigms of a more localized phenomenon relevant to stem shape, which nonetheless encompasses the whole Romanian branch (in the wider sense). At a stage prior to the division of this branch into the four dialect groups recognized today, stressed *e* and *o* underwent diphthongization to *ea oa* in certain phonological

contexts (cf. Loporcaro 2011: 128-30, with the references given there, for discussion of contrasting views on the precise nature of this change). This gave rise not only to the second conjugation infinitive ending *-eá*, which in diachronic terms can be taken as a variant of the stressed theme vowel *e*, but also to the morphophonological alternations *e/ea* and *o/oa* in many verbs: thus note the 3sg present *poáte*, *doárme* and 3sg/3pl subjunctive forms *poátă*, *doármă* alongside 1sg present *pot*, *dorm*.¹⁰⁵ Before an *e* in the following syllable, however, the *ea* which resulted from this process has since monophthongized to *e* again. This results in the divergence in stem shape between 3sg present *țése* and 3sg/pl subjunctive *țeásă*, corresponding to the opposition between present *-e* and subjunctive *-ă* in non-first conjugation verbs. But in the first conjugation this opposition is reversed, and so is the associated stem alternation: thus from *a întreba* ‘ask’ we see 3sg/pl present *întreábă*, 3sg/pl subjunctive *întrébe*. As will be seen, this pattern may continue to mark conjugational affiliation in cases where the evidence of the endings themselves is compromised.

¹⁰⁵ This is not an ‘automatic’, phonological matter: contrast 3sg/pl present *poártă* from inherited *a porta* ‘carry’, which shows the alternation, and 3sg/pl present *transpórtă* from borrowed *a transporta* ‘transport’, which does not (cf. Zafiu 2013: 28).

4.2 Heterocclisis as the indirect result of sound change: Maiden 2009

The phenomenon investigated by Maiden 2009 concerns the inflectional behaviour, in certain local dialects of Romanian, of the two verbs which in the standard variety appear as *a țese* ‘weave’ and *a coase* ‘sew’. These descend from the Latin third conjugation verbs *texō -ere* ‘weave’ and *cōnsuō -ere* ‘sew together’, and continue to inflect according to the third conjugation in standard Romanian too. I give the paradigm of *a coase* in Table 4.7 below: note that while *a coase* shows allomorphy between *cus-* (in the preterite, pluperfect and past participle) and *cos-* (elsewhere) which is not mirrored in the behaviour of *a țese* (Table 4.3), alternations following this pattern are commonly found among third conjugation verbs in Romanian, as mentioned above with regard to *a arde* in Table 4.4; in the standard language there is nothing anomalous about the behaviour of *a țese* and *a coase* as third conjugation verbs.

Table 4.7 The paradigm of *a coase* ‘sew’ in standard Romanian

Infinitive *coáse*
Gerund *cosánd*
Past participle *cusút*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>cos</i>		<i>cos</i>	<i>coseám</i>	<i>cusúi</i>	<i>cusúse</i>
2sg	<i>cóși</i>	<i>coáse</i>	<i>cóși</i>	<i>coseái</i>	<i>cusúși</i>	<i>cusúseși</i>
3sg	<i>coáse</i>		<i>coásă</i>	<i>coseá</i>	<i>cusú</i>	<i>cusúse</i>
1pl	<i>coásem</i>		<i>coásem</i>	<i>coseám</i>	<i>cusúrăm</i>	<i>cusúserăm</i>
2pl	<i>coáseți</i>	<i>coáseți</i>	<i>coáseți</i>	<i>coseáți</i>	<i>cusúrăți</i>	<i>cusúserăți</i>
3pl	<i>cos</i>		<i>coásă</i>	<i>coseáu</i>	<i>cusúra</i>	<i>cusúseră</i>

However, as documented by Maiden on the basis of published and unpublished dialect survey data from *NALR* (*Noul atlas lingvistic român pe regiuni* ‘The new Romanian linguistic atlas by regions’), a characteristic and widespread dialectal sound change has led, in many local varieties of Romanian, to distinctive behaviour

in the inflection of these two verbs: rather than belonging straightforwardly to the third conjugation in these varieties, they show heteroclisis.

The sound change in question affects the front vowels *i* /i/ and *e* /e/ and what can be called the ‘front diphthong’ *ea* /ɛa/. In the relevant areas, these segments have undergone centralization after the coronal fricatives *s* /s/, *z* /z/, *ș* /ʃ/, *j* /ʒ/, *ț* /ts/ (and /dz/ where this exists separately from /z/), giving the vowels *â* /â/, *ă* /ə/ and *a* /a/ respectively: cf. Caragiu-Marioțeanu (1975: 153) on the general development involved and e.g. Neagoe (1984: 244f.) and Marin and Marinescu (1984: 369) on its regional manifestations.

The relevance of this sound change purely to the phonetic shape of dialectal forms of *a coase* and *a țese* is evident. In those varieties where it occurred, it would automatically have operated upon all those parts of the paradigm in which the *s* of the root originally preceded an inflectional *e* or *ea* inherited as expected for third conjugation verbs. As can be seen from Tables 4.3 and 4.7, which show the historically underlying distribution of these sequences as retained by the standard language, the forms implicated make up a substantial amount of the paradigm. We would expect to see the results of this sound change not only throughout the imperfect (standard *coseam*, *țeseam* etc.), which would thus show centralized *-sam* etc., but also in the various forms historically containing the sequence *se*, including the infinitive and the 3sg present, which would produce *să*. Notably, by coincidence, in some of these cells this development would efface the surface distinction between the third conjugation morphology of these verbs and the regular morphology expected of verbs belonging to the *first* conjugation: taken in isolation, an imperfect in *-sam* or a 3sg present in *-să* could belong to either. However, as a phonological effect only, this need not in itself compromise the third conjugation identity of these

verbs; and alongside the ambiguous forms there would remain many which remained unambiguously affiliated with the third conjugation even after the sound change, such as the rhizotonic infinitive and 1/2pl present forms.

All these expected effects can indeed be identified in the paradigm of *a coase* in the northern Oltenian variety of Dobrița, provided in Table 4.8 (following Maiden’s Table 12, p. 71, based on the data provided by *NALROltenia* vol. 5), which will provide a starting point for our discussion of the type of heterocclisis at issue here. In this variety the centralization of front vowels has taken place, giving e.g. *măsá* ‘molar’ (cf. standard *măseá* < *maxilla* ‘jaw’) and its plural *măsále* (cf. standard *măsele*). It is thus unsurprising that, in this paradigm, alongside other slight phonological divergences from the standard, the effects of centralization after *s* are visible in the imperfect *cosám* etc (cf. standard *coseám*), as well as in every form continuing an earlier *se*, namely infinitive, 3sg present and 2sg imperative *cuasă* (cf. standard *coase* for all three), 1pl present and subjunctive *cuásăm* (cf. standard *coásem*), and 2pl present, imperative and subjunctive *cuásăț* (cf. standard *coáseți*). Analogous behaviour is shown by *a țese*. The sequences *se* and *sea* are apparently absent from the variety altogether, as a result of the sound change under discussion.

Table 4.8 The paradigm of *a coase* in Dobrița

Infinitive *cuásă*
 Gerund *cosán*
 Past participle *cusút*

	present	imperative	subjunctive	imperfect	preterite
1sg	<i>cos</i>		<i>cos</i>	<i>cosám</i>	<i>cusúí</i>
2sg	<i>cóși</i>	<i>cuásă</i>	<i>cóși</i>	<i>cosái</i>	<i>cusúși</i>
3sg	<i>cuásă</i>		<i>cuásă</i>	<i>cosá</i>	<i>cusú</i>
1pl	<i>cuásăm</i>		<i>cuásăm</i>	<i>cosám</i>	<i>cusúrăm</i>
2pl	<i>cuásăț</i>	<i>cuásăț</i>	<i>cuásăți</i>	<i>cosáți</i>	<i>cusúrăți</i>
3pl	<i>cuásă</i>		<i>cuásă</i>	<i>cosáu</i>	<i>cusúră</i>

Crucially, however, the conjugational behaviour seen in Table 4.8 cannot be fully explained on the basis of this diachronic process of centralization, or on the basis of a

resulting synchronic constraint ruling out *se* and *sea*. One form, at least, stands out, namely the 3pl present *cuásă*. The failure of the Dobrița variety to retain the expected form *cos* here (as is seen in the standard language, ultimately continuing Latin *consuunt*) cannot plausibly be treated as a question of sound change or synchronic phonology, especially as the form *cos* is found in the 1sg present and subjunctive of this verb as in the standard: there is simply no way to arrive at the attested form *cuásă* by phonetic means alone.

However, sense can be made of 3pl present *cuásă* if we recognize that a *morphological* development has taken place here. This is because although *cuásă* is unexpected in this cell, it is far from unique within the paradigm; in particular, this form is shared by the 3sg present, where we have seen that it arose by regular sound change. As mentioned above, syncretism between 3pl and 1sg forms in the present, as is displayed by *a coase* and *a țese* in standard Romanian, is a feature generally restricted to non-first conjugation verbs; meanwhile, syncretism between the two third person forms in the present, which are generally characterized by a final *-ă*, is a feature of first conjugation inflection, reflecting the phonetic merger between the reflexes of Latin *-at* and *-ant*. The remarkable 3pl present *cuásă*, syncretic with the 3sg, can thus be understood if we assume that it arose thanks to the adoption of distinctively first conjugation morphology in this particular cell.

The trigger for such unusual morphological behaviour is readily identifiable, given what was said above about the neutralizing effect of the sound changes undergone by certain forms of the verb. Evidently first conjugation *cuásă* did not emerge in isolation to replace the inherited third conjugation *cos*; rather, speakers already interpreted the paradigm as containing first conjugation morphology in certain cells where the distinction between the two conjugations had been effaced by

the centralization process. The appearance of 3pl *cuásă* in the paradigm in fact serves to demonstrate that this morphological reanalysis occurred. In particular, the adoption of this pattern of syncretism presupposes that 3sg present *cuásă* was treated as a first conjugation form at the time when the reshaping of the 3pl form took place. That is, speakers came to take the ending *-ă* ‘at face value’ as a marker of first conjugation membership, rather than as a phonological variant of third conjugation *-e* conditioned by the preceding fricative. At the same time, certain other forms in the paradigm must still be treated as belonging to the third conjugation: as mentioned above, rhizotonic forms such as 1pl present *cuásăm* show unmistakable third conjugation morphology. The paradigm of *a coase* in the variety of Dobrița is thus heteroclitic.

It may seem premature to identify heteroclitism in this variety, and make historical inferences about its emergence, simply on the basis of the single unexpected 3pl present form *cuásă* – although it should be said that no alternative explanation is readily available. However, the type of anomalous conjugational behaviour apparently shown by 3pl *cuásă* is very far from unique. As Maiden points out (p. 72, fn. 17), even from within the variety of Dobrița itself there is a further indication that this morphological reanalysis took place. This concerns the behaviour of *a țese* in the 3sg present indicative and (3sg/3pl) subjunctive. As was noted above, many Romanian verbs, including *a țese*, show an alternation between *e* and *ea* in the root in these two forms: phonetic in origin, this alternation sees *e* in the indicative and *ea* in the subjunctive (thus *țese* and *țeásă* respectively), in all conjugations except the first, where the reverse is found. That is, the root vowel and the ending are correlated, root *e* appearing together with final *e* and the diphthong appearing in the presence of final *-ă*. In Dobrița, then, one might expect to find indicative *țăsă* and

subjunctive *țásă*, representing ‘centralized’ versions of the regular third conjugation forms seen in the standard. But in fact the *reverse* is found. This suggests not only that the indicative form has been reinterpreted as belonging to the first conjugation, with a ‘genuine’ ending *-ă*, but also that, on this basis, in the subjunctive form speakers have secondarily applied the first conjugation ending *-e*, which has then undergone a further round of centralization. Meanwhile, there is no analogous alternation in Romanian which would allow us to detect the same diachronic developments in the paradigm of *a coase* in the variety of Dobrița; nonetheless, we may infer that they took place. Although 3pl *cuásă* is the only form in Table 4.8 which is obviously in need of an explanation beyond regular sound change, it is likely that the 3sg/3pl subjunctive form was also reshaped to pattern with the first conjugation as was its counterpart in the paradigm of *a țese*.

For these two verbs, conjugational behaviour of the kind exemplified by the Dobrița variety is found across northern Oltenia: evidently on the basis of a reanalysis of the conjugational affiliation of the ambiguous 3sg present form, both the 3pl present and the 3sg/3pl subjunctive have been remodelled as first conjugation forms amid paradigms which, however, still largely continue inherited third conjugation morphology (and meanwhile the 2sg imperative retains its syncretism with the 3sg present, as would be expected of a first conjugation verb). Notably, some varieties even preserve the newly adopted first conjugation subjunctive in *-e* without subjecting it to a second process of centralization, thus making the existence of heteroclisis in these verbs all the more apparent.¹⁰⁶

The emergence of this pattern, whereby the 3sg present, 3pl present and 3sg/3pl subjunctive all hang together with regard to their conjugational affiliation, is

¹⁰⁶ Varieties showing this behaviour include, for example, Bustuchin (point 918), Balș (point 934), Osica de Jos (point 986).

replicated again and again in the development of various heteroclitite phenomena arising in the history of the Romanian verb – although, as will be seen, these may differ in terms of both the initial trigger for the morphological change and the synchronic situation which eventually results. However, innovative first conjugation morphology in the paradigm of *a țese* and *a coase* which is inexplicable in terms of sound change may also be found outside these cells. Consider Table 4.9, showing the paradigm of *a coase* in the central Oltenian variety of Beloțu (point 970): Maiden notes that the conjugational behaviour seen in this variety is replicated in both *a coase* and *a țese* throughout central and southern Oltenia, meaning that in effect it constitutes a different regional norm from that discussed above for northern Oltenia.

Table 4.9 The paradigm of *a coase* in Beloțu (after Maiden 2009: 66f.)

Infinitive *cusá*
 Gerund *cusánd*
 Past participle *cusút*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>cos</i>		<i>cos</i>	<i>cusám</i>	<i>cusúi</i>	<i>cusúsăm</i>
2sg	<i>coș</i>	<i>coásă</i>	<i>coș</i>	<i>cusái</i>	<i>cusúși</i>	<i>cusúsăși</i>
3sg	<i>coásă</i>		<i>coáse</i>	<i>cusá</i>	<i>cusú</i>	<i>cusúsă</i>
1pl	<i>cusám</i>		<i>cusám</i>	<i>cusám</i>	<i>cusúrăm</i>	<i>cusúsăm</i>
2pl	<i>cusát</i>	<i>cusát</i>	<i>cusát</i>	<i>cusát</i>	<i>cusúráț</i>	<i>cusúsăț</i>
3pl	<i>coásă</i>		<i>coáse</i>	<i>cusáu</i>	<i>cusúrá</i>	<i>cusúsă</i>

The forms given in bold are those which pattern distinctively with the first conjugation; in the preterite, pluperfect and past participle the verb still retains its original third conjugation morphology, while the gerund and 1sg and 2sg present/subjunctive forms cannot be assigned unambiguously to either conjugation. The interesting forms for our purposes here are the 1pl present and subjunctive, the 2pl present, imperative and subjunctive, and the infinitive: unlike in the northern Oltenian pattern, these all show first conjugation morphology which cannot have arisen as a direct result of the centralization process – along with all those parts of the paradigm which already did so in Dobrița. The fact that the third person

subjunctive forms are marked by *-e*, without centralization after *s*, indicates that the imperfect in *-ám* (as opposed to *-eám*) is a genuine first conjugation feature.

The diachronic implication of this paradigm is clear: the state of affairs exemplified here represents a more advanced stage in the intra-paradigmatic spread of first conjugation morphology than that seen above for northern Oltenian. After the initial reinterpretation of certain ambiguous forms in the paradigm (in particular the 3sg present) as showing first rather than the third conjugation morphology, this first conjugation affiliation was first of all extended to the other third person present and subjunctive forms along with the second person imperative, giving the situation still seen in northern Oltenia and exemplified in Table 4.8; the behaviour seen in central and southern Oltenia results from a subsequent further expansion of the ‘domain’ of first conjugation morphology within the paradigm, which now covers all cells (to which the conjugational distinction is relevant) outside the preterite, pluperfect and past participle.

The patterns of heteroclisis presented for dialectal *a coase* and *a țese* thus result from the analogical spread of first conjugation morphology through what were originally regular third conjugation paradigms. Equally important, however, is the fact that this analogical spread only progressed so far through the paradigm and no further; were this not the case, the outcome would naturally be a verb showing no heteroclisis but instead belonging entirely to the first conjugation, an outcome which is in fact attested at precisely one survey point in Oltenia, namely Castranova (point 987).¹⁰⁷ But given that the extension of first conjugation morphology is generally halted in these varieties before it has covered the whole paradigm, it is interesting to observe the systematicity of the distributions found: either the third persons singular

¹⁰⁷ Where, however, an optional past participle in *-út* also survives.

and plural of the present and the subjunctive (together with the 2sg imperative) take on the innovative affiliation with the first conjugation, or it is generalized to the whole paradigm with the exception of the preterite, pluperfect and past participle.

As Maiden points out (2009: 77f.), it is not clear that these particular patterns of morphological behaviour can be addressed by appealing to functional features or aspects of the morphosyntactic structure of the paradigm. With regard to the more advanced stage in the development of heteroclisis seen in central and southern Oltenia, it may appear possible to account for the consensus of the preterite, pluperfect and past participle in terms of perfectivity (rather than past tense, given the divergent behaviour of the imperfect), but there is no evidence that the Romanian verb is still articulated by an morphosyntactic distinction of aspect as was the Latin verb, with its paired *perfectum* and *imperfectum* series characterizing the whole of the finite paradigm (though note the difficulty already found in pinning down the nature of this distinction even for Latin: section 3.1.1). However this may be, the initial stages of the spread of heteroclisis are still harder to explain in functional terms. Following a reanalysis of conjugation class membership in the 3sg present, it may seem unremarkable if an early analogical extension of this reanalysis should implicate the 3pl present; given the lack of any functional link between the Romanian present and subjunctive, however, there is no reason for the 3sg/pl subjunctive to be systematically affected at the same time, while the remaining present tense forms and all the remaining third person forms in the paradigm are, equally systematically, left unaffected.

On the basis of what was said in the overview of Romanian verb inflection above, however, it can be seen that these groups of cells are entirely coherent in purely morphological terms. All forms of the perfect, all forms of the pluperfect, and

the past participle already diverge from the rest of the paradigm in terms of stem shape in the second and third conjugations, usually all sharing a single distinctive stem themselves. Meanwhile, the third persons present and subjunctive (together with the 2sg imperative) all pattern together elsewhere in Romanian verb morphology, in that they make use of the distinctive N-pattern stem – the only other relevant forms (in the majority of verbs showing N-pattern alternation) being the 1sg and 2sg present/subjunctive, which show no conjugationally specific marking at all in *a țese* and *a coase* whether or not these show any signs of heteroclisis. In behaving as they do with regard to the analogical extension of first conjugation morphology, speakers can thus be seen to show sensitivity to the morphomic properties underlying the makeup of the verb paradigm in general.

Accordingly, there may be no reason to suppose, with Stump 2006, that heteroclisis is fundamentally a phenomenon constrained (directly or indirectly) by properties of a language's morphosyntactic structure, since instances of heteroclisis can be found, as here, which instead directly reflect purely morphological principles of paradigmatic organization. On the contrary, Maiden proposes (2009: 81-85) that *all* heteroclisis may turn out to be morphologically conditioned in that it follows existing patterns of stem alternation, as in these Romanian cases – and thus that any apparent instance of morphosyntactic conditioning in heteroclisis in fact echoes an pattern of stem alternation already correlated with the morphosyntactic split.

The inflectional behaviour of the Latin *capiō* type, treated in 3.2 above, represents a counterexample to Maiden's proposal, since the synchronic distribution of third and fourth conjugation morphology over the *imperfectum* cells of *capiō* and the other verbs of this type does not echo a pattern of stem alternation which was already found elsewhere in the Latin verb system. However, as we have seen, the

emergence of heteroclisis in the *capiō* type is somewhat different from that seen in the verbs treated here: the pattern seen in *capiō* does not result from the *redistribution* of conjugational morphology across a paradigm, but appears to be the original distribution produced by the sound change which gave rise to heteroclisis in the first place. The existence of such cases of heteroclisis suggests that Maiden's proposal should rather be treated as a diachronic principle concerning the adoption of historically 'intrusive' conjugational morphology into a paradigm. In the remainder of this chapter, we will see that the relevance of morphologically coherent zones of the paradigm to the patterns taken up by heteroclisis is widely observable in the development of Romanian – even in cases of heteroclisis which have not been triggered, as in *a coase* and *a țese*, by the operation of neutralizing sound change on existing regular forms. I turn to two such cases now, which exemplify different ways in which a lexeme's paradigm can demonstrate the influence of two distinct conjugations without any contribution from neutralizing sound change: in both cases the distribution of conjugational behaviour over the paradigm is found to echo identifiable morphomic patterns visible elsewhere in Romanian. I then (section 4.4) turn back to cases historically dependent on sound change, and thus similar to what has just been seen for *a coase* and *a țese*, which are not attested only in the Romanian dialects but have come to characterize the standard language as well.

4.3 Beyond *a coase* and *a țese*

Most varieties of Romanian, including the standard, possess a verb meaning ‘hiccup’ which continues an etymon **suggluttiō -āre*, generally taken to represent a version of *singultō -āre* ‘gasp, hiccup, sob’ with contamination from *glut(t)iō -īre* ‘swallow’ (cf. Spanish *sollozar*, Portuguese *soluçar*: see *REW* s.v. *singultāre*). In standard Romanian this verb takes the form *a sughita* and inflects as a regular unaugmented first conjugation verb (see e.g. Uricaru and Goga 1997):

Table 4.10 The paradigm of *a sughita* ‘hiccup’ in standard Romanian

Infinitive *sughită*
 Gerund *sughitănd*
 Past participle *sughitat*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>sughit</i>		<i>sughit</i>	<i>sughitam</i>	<i>sughitai</i>	<i>sughitasem</i>
2sg	<i>sughiti</i>	<i>sughită</i>	<i>sughiti</i>	<i>sughitai</i>	<i>sughitashi</i>	<i>sughitaseși</i>
3sg	<i>sughită</i>		<i>sughite</i>	<i>sughita</i>	<i>sughita</i>	<i>sughitase</i>
1pl	<i>sughităm</i>		<i>sughităm</i>	<i>sughitam</i>	<i>sughitarăm</i>	<i>sughitaserăm</i>
2pl	<i>sughitați</i>	<i>sughitați</i>	<i>sughitați</i>	<i>sughitați</i>	<i>sughitarăți</i>	<i>sughitaserăți</i>
3pl	<i>sughită</i>		<i>sughite</i>	<i>sughitau</i>	<i>sughitară</i>	<i>sughitaseră</i>

However, Lombard (1955: 595), citing Jordan (1948: 127), notes the existence ‘inconnue des dictionnaires’ of a rival infinitive *sughiti*, evidently belonging to the fourth conjugation. In an earlier article Jordan in fact observes (1935: 62) that fourth conjugation *sughiti* is widespread (indeed the only guise in which the verb is found) in the Romanian of Moldova, while Pop (1952: 228) records an augmented form *sughitésc*, demonstrating the admission of this lexeme, for some speakers, into the productive subclass of the fourth conjugation.

As Lombard points out, the existence of fourth conjugation forms of this verb reflects the influence of the inflection of the more commonplace *a înghiți* ‘swallow’, itself continuing prefixed *ingluttiō -īre* with the same meaning. However, although in some localities there is no sign of fourth conjugation morphology associated with

a sughița, while others, by contrast, retain no trace suggesting that it ever belonged to the first conjugation, the detailed dialectological evidence available to us shows that the adoption of innovative fourth conjugation variants in the paradigm of *a sughița* is not uniformly an all-or-nothing affair. Instead, we frequently find heteroclite paradigms reported for local varieties, which make use of both first and fourth conjugation morphology assigned to different cells. Notably, however, the paradigmatic distribution of forms representing the two conjugations does not simply differ arbitrarily across the records taken at the various survey points; rather, systematic preferences are detectible in the way speakers allot conjugational morphology over the paradigm.

This can be illustrated on the basis of data from Banat, drawn from survey maps in *NALRBanat* vol. 4 (Beltechi et al. 2005).¹⁰⁸ The Banat region lies at the far (south-)western edge of Romania, and unlike the Moldova region, lying at the other geographical extreme of the Daco-Romanian dialect area, it has not been associated in the scholarly literature with the development of fourth conjugation morphology in *a sughița* so prevalent in the latter. And although the fine-grained dialectological survey whose findings are presented in *NALRBanat* collected responses on only the present (and sometimes the 2pl imperative), subjunctive and infinitive forms of the verb rather than its entire inflectional paradigm, the data which are available strongly bear out the impression that this region remains conservative with regard to the adoption of variant fourth conjugation forms of this verb. Thus at more than sixty of the hundred survey points used in the region we find only first conjugation forms in the present, subjunctive and infinitive, which diverge little (and most usually not at

¹⁰⁸ The maps in question are numbers 624-28, which provide a geographical synthesis of the responses to questions 205 (present), 206 (subjunctive) and 207 (singular and plural prohibitive, furnishing the infinitive and 2pl imperative respectively) from the master questionnaire which serves as the basis for all regional editions of *NALR*.

all) from those found in the standard paradigm presented just above – with the exception that the centralization process seen above for *a coase* and *a țese* has also applied in these varieties, giving third person subjunctive *sughîță* for standard *sughîte*.

At the same time, however, a fully fourth conjugation paradigm is recorded for this verb at seven survey points in the region,¹⁰⁹ as for example at Plavișevița (point 4):

Table 4.11 The paradigm of *a sughîța* in Plavișevița (Banat)

	infinitive	<i>sughîți</i>	
	present	imperative	subjunctive
1sg	<i>sughît</i>		<i>sughît</i>
2sg	<i>sughîț</i>	[not given]	<i>sughîț</i>
3sg	<i>sughîce</i>		<i>sughîță</i>
1pl	<i>sughîțîm</i>		<i>sughîțîm</i>
2pl	<i>sughîțîți</i>	<i>sughîțîți</i>	<i>sughîțîți</i>
3pl	<i>sughît</i>		<i>sughîță</i>

Several features mark this paradigm out as having abandoned the first conjugation in favour of the fourth. In particular, the vowel *î* found in the 1pl and 2pl forms and the infinitive results from the centralization of the theme vowel *i* seen for these forms in the fourth conjugation in standard Romanian; meanwhile, final *-e* in the 3sg present, final *-ă* in the 3sg/3pl subjunctive and the absence of an inflectional ending in the 3pl present are all non-first conjugation features. Lastly, the alternation between stem-final *ț* and *t* seen in this paradigm (where the *ce* [ce] of the 3sg present is the regular reflex of *te* across Banat, cf. the forms given in map 32 (*frunte* ‘forehead’) in Saramandu 2005) is unknown in first conjugation *a sughîța*, but is a feature of practically all fourth conjugation verbs with an infinitive in *-ți*, including *a înghiți* (Lombard 1955: 595-99). Note that the 2sg present/subjunctive form in *-ț* (standard *-ți* with asyllabic vowel) is a constant over both types of conjugational behaviour.

¹⁰⁹ The survey points concerned are 4 (Plavișevița), 13 (Lescovița), 19 (Răcășdia), 28 (Secășeni), 47 (Pădureni), 56 (Comloșu Mare), and 64 (Igrîș).

The influence of fourth conjugation morphology on the inflection of *a sughîta* in Banat is not, however, primarily reflected by the existence of consistently fourth conjugation paradigms such as that seen in the variety of Plavişeviţa (point 4). Much more numerous are survey points at which fourth conjugation morphology is found only in certain of the inflectional forms investigated, but not all. But in such cases one specific distribution of the innovative conjugational morphology predominates,¹¹⁰ as illustrated here by Table 4.12, which gives the forms recorded for *a sughîta* at Cornereva (point 26):

Table 4.12 The paradigm of *a sughîta* in Cornereva (Banat)

	infinitive	<i>sughîta</i>	
	present	imperative	subjunctive
1sg	<i>sughît</i>		<i>sughît</i>
2sg	<i>sughîţ</i>	[not given]	<i>sughîţ</i>
3sg	<i>sughîce</i>		<i>sughîă</i>
1pl	<i>sughîăm</i>		<i>sughîăm</i>
2pl	<i>sughîăţ</i>	<i>sughîăţ</i>	<i>sughîăţ</i>
3pl	<i>sughît</i>		<i>sughîă</i>

The forms given in bold are those which do not reflect the regular behaviour of first conjugation *a sughîta*, instead showing the influence of the fourth conjugation on the paradigm. As can be seen, innovation is found in the 1sg, 3sg and 3pl forms, while conservative forms are retained in the 1pl and 2pl and in the infinitive; the 2sg forms do not reflect any conjugational affiliation.

No other distribution of first and fourth conjugation morphology offered by informants for the relevant varieties can be compared to this in terms of breadth of attestation over the region. Indeed, most other patterns which could be listed are unique to single survey points and are often recognizably ‘unsystematic’ on other

¹¹⁰ This particular pattern of heteroclite behaviour, or others hardly diverging from it, can be seen at points 1 (Jupalnic), 2 (Ieşelniţa), 6 (Sicheviţa), 9 (Cărbunari), 16 (Pecinişca), 20 (Borlovenii Vechi), 23 (Vărădia), 24 (Brădişoru de Jos), 26 (Cornereva), 27 (Gârlişte), 29 (Văliug), 31 (Bucuşniţa), 32 (Doman), 35 (Lăţunaş), 37 (Ferendia), 74 (Constantin Daicoviciu), 75 (Zgribeşti) and 79 (Jdioara).

grounds. At Lugojel (point 81) alone, for example, a fourth conjugation feature (final *t* in place of conservative *Ń*) is recorded only for the 1sg *subjunctive* form *sughit* in an otherwise fully regular first conjugation paradigm; but note that in the dialect of Banat, as throughout Romanian, the formal distinction between 1sg present and 1sg subjunctive has been entirely lost, meaning that the purported uniqueness of the 1sg subjunctive in Lugojel cannot be genuine but must be an artifact of the elicitation process entailed by the dialect survey.

In contrast, the fact that the pattern exemplified in Table 4.12 is widely replicated over the region vouches for its morphological reality. Indeed, this pattern is significantly better represented in Banat than that which sees fourth conjugation morphology in all the recorded forms. In this light, it is notable that the morphological division which emerges from the data directly parallels that identified in the heteroclite behaviour of *a țese* and *a coase* in the previous section, in which 3sg and 3pl forms in the indicative and subjunctive (the 1sg forms being conjugationally ambiguous in that case) likewise patterned together against the remaining plural forms and the infinitive with regard to conjugation class.

Again, then, we see evidence here for the importance of the N-pattern in articulating conjugational affiliation in addition to stem allomorphy over the paradigm of verbs in Romanian. This finding is particularly striking in light of the fact that the motivation for conjugational innovation in *a sughita* is utterly unlike that seen for the above verbs: in the present case, rather than morphological ambiguity on the basis of phonological neutralization, it has come about thanks to the similarity between two historically related lexemes in terms of both stem shape and meaning. It is also notable that, in this case, the innovative characteristic displayed by the N-pattern cells is a conjugational shift *away* from the first conjugation, rather than

towards it as in *a țese* and *a coase*; if a generalization is to be made, it is not that the N-pattern cells are predisposed towards first conjugation morphology in some sense, but that they represent a coherent locus of early conjugational innovation in the paradigm.

These dialect data on the inflectional morphology of *a sughița* show that, even in a region of Romania in which the presence of innovative fourth conjugation forms of the verb has not generally been noted, the incursion of these forms into the paradigm due purely to the narrow influence of one lexeme on another has been sufficient to give rise to a widespread pattern of heteroclisis – which again reflects sensitivity to the N-pattern whose importance was seen earlier for the behaviour of *a țese* and *a coase*. In what follows we will see that, in individual cases, verbs can develop heteroclisis as the result of such analogical influences even where the lexical links involved may be less obvious than for *a sughița* with respect to *a înghiți*. Again, existing morphological patterns are revealed to play a part in the process.

An appreciation of the interest held by the next set of dialect data for the study of heteroclisis in Romanian requires familiarity with another morphomic pattern which features prominently in the history of the Romance verb, which is labelled (after Maiden 2011a) the ‘L-pattern’. The existence of this pattern in Romanian, as it concerns the types of verb stem involved in the cases treated here, originally came about as the result of a cluster of related sound changes which began to operate in Late Latin and the effect that these had on stem shape across the paradigm of certain groups of verbs. At issue is the palatalization/affrication of certain consonants immediately followed by the glide yod [j], which itself arose from an unstressed *e* or *i* in prevocalic position. Given the inflectional morphology of Latin as presented above (sections 3.1.1 and 3.2.1), it can be seen that yod from this

source emerged automatically as a characteristic of the 1sg present indicative, and the whole present subjunctive, of all verbs belonging to the second and fourth conjugations as well as verbs of the *capiō* type, along with the 3pl present indicative and the gerund of fourth conjugation and *capiō*-type verbs – a development which I illustrate in Table 4.13 for the fourth conjugation verb *saliō*, the ancestor of Romanian *a sări* ‘jump’:

Table 4.13 Partial paradigm of *saliō* ‘jump’ in Latin, and postulated forms after the development of yod

gerund <i>saliendō</i>			gerund <i>sa'ljendo</i>		
	pres. ind.	imperative	pres. subj.		
1sg	<i>saliō</i>		<i>saliam</i>	'saljo	'saljam
2sg	<i>salīs</i>	<i>salī</i>	<i>saliās</i>	'sales	'saljas
3sg	<i>salit</i>		<i>saliat</i>	'salet	'saljat
1pl	<i>salīmus</i>		<i>saliāmus</i>	sa'limus	sa'ljamus
2pl	<i>salītis</i>	<i>salīte</i>	<i>saliātis</i>	sa'lites	sa'ljatis
3pl	<i>saliunt</i>		<i>saliant</i>	'sal[j]unt	'saljant

The resulting clusters of consonant + yod underwent sound changes of various kinds across the Romance varieties (cf. Lausberg 1972: §§387-95, 451-78; Loporcaro 2011: 143-48), which led to the existence of regular alternations in stem form across the cells of the paradigm of the verbs involved. The upshot of this process for Romanian in particular can again be illustrated with the reconstructed reflexes of *saliō*. Table 4.14 displays the paradigm as postulated at a stage after the fusion of the cluster /l + j/ to give palatal /ʎ/, which thus alternated with /l/ over the paradigm. The forms comprising the L-pattern are given in bold.

Table 4.14 Illustration of the L-pattern (postulated) after fusion of the cluster /l + j/

gerund	sa'ʎendo		
	pres. ind.	imperative	pres. subj.
1sg	'saʎo		'saʎam
2sg	'sales	'sali	'saʎas
3sg	'salet		'saʎat
1pl	sa'limus		sa'ʎamus
2pl	sa'lites	sa'lite	sa'ʎatis
3pl	'salunt		'saʎant

The conventional label of ‘L-pattern’ for this pattern of alternation refers to the ‘right-angled’ distribution of the alternant reflecting the presence of Late Latin yod, which at the relevant historical stage was found throughout the present subjunctive but only in the 1sg of the present indicative. Note that, as mentioned above in section 4.1 (and symbolized by the form ‘sal[j]unt in Table 4.13), at some point early in the history of Romanian all non-first conjugation verbs generalized the 3pl present indicative ending *-unt* (as opposed to *-ent* or *-iunt*) which originally characterized the third conjugation in particular: as a result, the 3pl present indicative came to possess the alternant /l/ rather than phonologically expected /k/. On the other hand, the ‘yod alternant’ was generalized in the gerund to all verbs showing a stem alternation resulting from Late Latin yod, even if etymologically the gerund could not have acquired a yod by regular sound change. Thus among the numerous other verbs displaying the L-pattern for the same historical reason were the reflexes of *(re)maneō* and *videō*, which we can reconstruct as showing the alternations /n/ vs. /ɲ/ and /d/ vs. /dʒ/ corresponding to /l/ vs. /k/ in Table 4.14; in the gerund we can reconstruct these as displaying the unetymological alternants /ɲ/ and /dʒ/ respectively.

The eventual shape taken by the Romanian L-pattern is illustrated in Table 4.15, which displays the relevant forms of *a sări*. As /k/ gives /j/ in Daco-Romanian, while intervocalic /l/ > /r/, the alternation involved now features /r/ vs /j/ as a matter of regular sound change. But the distribution of the alternants themselves has been affected by the collapse, in the first and second persons, of the formal distinction between present indicative and subjunctive in favour of the indicative. The L-pattern forms, displaying the distinctive ‘yod alternant’, are thus now the 1sg present/subjunctive, the 3sg/3pl subjunctive, and the gerund. Note that the form *săie* < *salia(n)t* displays the result of the sound change /ə/ > /e/ after palatals (the effects

of this change still constitute a phonological constraint in Romanian, where schwa may not appear after a palatal, cf. Zafiu 2013: 28); as a result, in the 3sg the L-pattern alternation is all that marks the distinction between present and subjunctive.

Table 4.15 Partial paradigm of *a sări* illustrating stem alternation in L-pattern cells

gerund	<i>săind</i> [sə ^h iind]		
	present	imperative	subjunctive
1sg	<i>sai</i>		<i>sai</i>
2sg	<i>sári</i>	<i>sári</i>	<i>sári</i>
3sg	<i>sáre</i>		<i>súie</i>
1pl	<i>sárim</i>		<i>sárim</i>
2pl	<i>sáriți</i>	<i>sáriți</i>	<i>sáriți</i>
3pl	<i>sar</i>		<i>súie</i>

The forms given in Table 4.15, which represent the continuation of those in Table 4.14 given the general Romanian developments just mentioned, are not in fact those preferred for *a sări* in the modern standard language. In the standard, as widely across Daco-Romanian, the L-pattern as it is exemplified here has been subject to complex processes of levelling in recent centuries,¹¹¹ and in this particular verb all the forms in /j/ are now deprecated (in favour of 1sg present/subjunctive *sar*, 3sg/pl subjunctive *sară* etc.). Nonetheless, the L-pattern survives in many varieties, and in many verbs belonging to the standard, more successfully than it does in standard *a sări*, thus providing the conditions for the heteroclite phenomena I now present, discussed in the context of the L-pattern by Maiden (2011: 83f.).

The first of these, which is attested in a group of varieties in central Crișana (north-west Romania) as recorded in unpublished material from *NALRCrișana* (Stan and Urișescu 2000-), sees the involvement of the L-pattern alternation as it applies to the verb *a sări* itself. This involvement is somewhat cryptic, however, as the forms showing heteroclitism themselves belong to a different verb, namely *a zbura* (< **ex-volāre*), which in the standard language has the meaning ‘fly’. In this area, *a*

¹¹¹ This process is treated in detail in Maiden 2011, where wide differences of opinion are noted (p. 72) with regard to the date when this levelling began to operate.

zbura means not only ‘fly’ but also ‘jump’; its inflected forms are accordingly given, either alongside or (more commonly) instead of forms of *a sări*, at several survey points in response to questions 1631 and 1632 of the dialect survey, intended to elicit the present and subjunctive respectively of the latter verb. In the local varieties of interest to us,¹¹² it emerges that some forms of the paradigm of *a zbura* do not behave regularly according to the first conjugation; instead, stem-final yod is liable to appear in the L-pattern cells of 1sg present/subjunctive and 3sg/pl subjunctive, alternating with the *r* which would be expected to occur in all cells except the 2sg.¹¹³ Table 4.16 gives the forms recorded for Bălan (point 184), presented in a simplifying orthography in line with that of standard Romanian.

Table 4.16 Partial paradigm of *a zbura*, in the meaning ‘jump’, in Bălan (Crișana)

	present	subjunctive
1sg	<i>zboi</i>	<i>zboi</i>
2sg	<i>zboi</i>	<i>zboi</i>
3sg	<i>zbóră</i>	<i>zboie</i>
1pl	<i>zburám</i>	<i>zburám</i>
2pl	<i>zburáț</i>	<i>zburáț</i>
3pl	<i>zbóră</i>	<i>zboie</i>

Inflectional behaviour of this kind must be explained as the result of the analogical influence of *a sări* on its synonym in the relevant varieties. But this influence is not expressed equally across the paradigm – instead, its effects are restricted to the L-pattern cells, which take on the stem alternant shown in the corresponding forms of *a sări*; thinkable forms such as 3sg present **zboie* (after *sáre*) are not found.

This highly localized behaviour, whose existence has only been detectible thanks to the detailed and copious nature of the dialect survey underlying *NALR*,

¹¹² Stupini (point 169), Meseșenii (171), Cehal (177), Bocșa (179), Moigrad (182), Bălan (184), Nadiș (186), Stâna (189), Asuaju de Sus (199).

¹¹³ 2sg *zboi* (for standard *zbori*) is the expected form in this variety, and does not reflect any external conjugational influence: like all 2sg present/subjunctive forms (except those bearing an augment), its morphology is neutral with regard to conjugation class. The same applies for *a (a)mâna* and *a rămânea* in Table 4.17.

finds a parallel, however, in a more widespread phenomenon whose incongruous nature against the backdrop of the verb system as a whole has been noted several times in the literature (see e.g. Lombard 1954: 270-74; Orza 1979: 74f.; Zamfir 2005: 462; Maiden 2011a: 83). The two verbs *a mâna* ‘drive (cattle)’ and *a amâna* ‘postpone’ behave for the most part as regular verbs of the first conjugation (they continue *minor -ârî* ‘threaten’ and a derivative of *mâne* ‘in the morning’ respectively); however, many works of reference record forms of the 1sg present/subjunctive, 3sg/3pl subjunctive, and gerund (variously described as regional, optional, or even fully standard)¹¹⁴ in which the *n* of the root is given up in favour of *yod*.

Etymologically such forms are unwarranted, as L-pattern alternation is not an inherited feature of the first conjugation. The motivation for this behaviour is clear, however: the alternation is based on the example of the non-first conjugation verbs *a mânea* and more common *a rămânea* ‘stay’ (< *(re)maneō -ēre*; also found as third conjugation *a (ră)mâne*), which, as shown in Table 4.17, display this alternation as would be expected. The link here is not a semantic one¹¹⁵ but evidently concerns the shape of the root *mân-* itself.

Table 4.17 Partial paradigms of *a (a)mâna* and *a rămânea*

	gerund <i>(a)mâind</i> / <i>(a)mânând</i>		gerund <i>rămâind</i> / <i>rămânând</i>	
	present	subjunctive	present	subjunctive
1sg	<i>(a)mâi</i> / <i>(a)mân</i>	<i>(a)mâi</i> / <i>(a)mân</i>	<i>rămâi</i> / <i>rămân</i>	<i>rămâi</i> / <i>rămân</i>
2sg	<i>(a)mâi</i>	<i>(a)mâi</i>	<i>rămâi</i>	<i>rămâi</i>
3sg	<i>(a)mână</i>	<i>(a)mâie</i> / <i>(a)mâne</i>	<i>rămâne</i>	<i>rămâie</i> / <i>rămână</i>
1pl	<i>(a)mânăm</i>	<i>(a)mânăm</i>	<i>rămâném</i>	<i>rămâném</i>
2pl	<i>(a)mânăți</i>	<i>(a)mânăți</i>	<i>rămâneți</i>	<i>rămâneți</i>
3pl	<i>(a)mână</i>	<i>(a)mâie</i> / <i>(a)mâne</i>	<i>rămân</i>	<i>rămâie</i> / <i>rămână</i>

¹¹⁴ For example, see Lombard (1954: 271) for a table listing the varying judgements and preferences of different authorities on the distribution of ‘iotacized’ and ‘non-iotacized’ forms for these two verbs.

¹¹⁵ It could be argued that there is a semantic link between the notions of ‘stay’ and ‘postpone’. However, this still leaves *a mâna* ‘drive’ unaccounted for.

Maiden 2011, which discusses these cases primarily with regard to demonstrating the morphological reality of the L-pattern, raises but is somewhat non-committal on the question whether they should be considered to embody heteroclisis. In light of what has gone before, however, I see no reason not to identify heteroclisis in these patterns of conjugational behaviour. Just as it can be recognized on the basis of general traits of the conjugational system that a form such as 3sg present *sughiće* (for *sughíte*) does not 'belong' in the same paradigm as 1pl present *sughiţám*, it can be recognized that the alternation between *mâ/n/-* and *mâ/j/-* seen here represents the incursion of non-first conjugation behaviour into a first conjugation paradigm.

4.4 Standard Romanian phenomena

I turn first to a morphological phenomenon which, with regard to its diachronic rationale, echoes what was seen for *a țese* and *a coase* earlier in this chapter, as here again a specific type of conjugational shift can be seen to emerge as the indirect result of a phonological process of post-consonantal vowel centralization that itself operates blind to the morphology. However, for reasons which will become clear, the synchronic outcome of this change is somewhat different from that seen above: it is rather less evident that the behaviour of the resulting paradigm constitutes heteroclisis.

In the case treated here, centralization took place following what can be taken, for historical consistency and ease of reference, as a geminate /rr/ as opposed to single /r/ – the true phonetic distinction being between a trill and a tap, which have since fallen together in most varieties of Romanian, including the standard.¹¹⁶ The historical centralization of front vowels in this context is reflected across the Daco-Romanian varieties and may even be an ancestral feature inherited in common by both Daco-Romanian and Aromanian (Rosetti 1968: 523, following Sala 1964), although this is not certain (Schulte 2005: 379). Classic examples of the development in standard Romanian are seen in *râu* ‘river’ < *rīvus* ‘stream’ and *rău* ‘bad’ < *reus* ‘guilty’, in which the centralization of the vowels reflects the development of a trilled initial consonant.

In the conjugation of verbs undergoing it, this sound change led to morphological reanalysis within the paradigm in precisely the same way seen for dialectal *a țese* and *a coase* above, as shown by the behaviour of the (originally third conjugation) reflex of Latin *currō -ere* ‘run’. As collated by Zamfir (2005: 115-20),

¹¹⁶ The distinction survives only in Maramureș and northern Ardeal, where a trill is preserved in words such as *rău* ‘bad’ (Rosetti 1968: 522).

attestations from the Romanian of the sixteenth and seventeenth centuries demonstrate the incursion of first conjugation morphology into its paradigm, e.g. the unambiguously first conjugation 3pl present *cúrá*¹¹⁷ and infinitive *curá* (alongside the phonologically regular infinitive *cúră*); this development must have come about as the result of the conjugational ambiguity of phonologically regular forms after centralization, such as 3sg present *cúră* (< **curre* < *currit*).

However, I am interested here in the morphological upshot of the same sound change as it affected verbs previously belonging to the *fourth* conjugation. As representative of the small group of verbs in question here, numbering around sixteen in total, I take *a omorî* ‘kill’: the effect of the historical process of centralization is evident on the final vowel of this infinitive form. This is a loanword from Slavic (< *umoriti* ‘kill’), as indeed are practically all the verbs of this type, strongly suggesting that /rr/, while existing in native Romanian words, was treated at some point as a stereotypical marker of non-native vocabulary and generalized as such (Schulte 2005); in this group only *urî* ‘hate’ < **horrîre* (cf. Latin *horreō* ‘bristle’) is inherited from Latin.

One clear signal of the historical relationship between these verbs and the Romanian fourth conjugation is the fact that the majority of verbs of this type, including *urî*, make use of an augment with allomorphs *-ăsc/-asc/-ăşt-* which is found in those parts of the paradigm where an augmented fourth conjugation verb displays *-esc/-easc/-eşt-*, e.g. 1sg present *urăsc* ‘I hate’, cf. 1sg *citesc* ‘I read’: the augment in these verbs clearly corresponds exactly to that displayed in the fourth conjugation, with the sole distinction that it displays the effects of the centralization process resulting from the former presence of /rr/. Indeed, the shape taken by any

¹¹⁷ Forms of this antiquity are transliterations from Cyrillic (the accent is added here for exposition). Note that the orthography did not reliably express the difference between the trill and the tap.

inflected form of these verbs continues to echo precisely the corresponding regular fourth conjugation form.

In verbs such as *a omorî*, however, which do not¹¹⁸ make use of this augment – thus corresponding historically to the type exemplified above by *a dormi* – we see, yet again, that phonological centralization has given rise to a conjugational shift, in which first conjugation morphology is introduced into the paradigm. As with the reflex of *currō*, the effect of this morphological development is already visible in texts from the sixteenth century: thus alongside 3sg present *omoară*, which is the phonologically expected outcome of a fourth conjugation form with the shape **umorre* (or similar) but happens to resemble a first conjugation form, we find not only the ‘regular’ 3pl present *omor* but also *omoară*, showing the distinctive 3sg/3pl syncretism characteristic of the first conjugation. In the modern standard this incursion of first conjugation morphology has won out altogether, leaving *omoară* as the only existing 3pl form; first conjugation morphology has also been generalized to the 3sg/3pl subjunctive, which appears as *omoare* – a form which evidently does not result directly from centralization, but again embodies the incursion of first conjugation morphology into the N-pattern cells. Maiden (2009: 79) further points out that there is reason to think that the 2sg imperative of such verbs has also become associated with the morphology of the first conjugation in particular: given the observation made above that, outside the first conjugation, this form is syncretic with the 3sg present only for transitive verbs, it is notable that intransitive *a coborî* ‘go

¹¹⁸ This is true of the modern language. However, Lombard (1955: 819) notes that this verb could take the augment at an earlier stage. This is corroborated by Zamfir (2005: 212), but her data are nonetheless sufficient to show the incursion of the first conjugation morphology into the unaugmented 3pl present.

down'¹¹⁹ nonetheless has the 2sg imperative *coboară* (= 3sg present; cf. 2sg present *cobori*), unambiguously signalling its affiliation with the first conjugation.

The relevance of the centralized 3sg present *omoară* (etc.) for the development of first conjugation morphology in the present and subjunctive of *a omori* and the few verbs that pattern with it¹²⁰ was noticed by Morariu (1924: 68f.), who suggested the analogy of first conjugation verbs of the same shape such as *a zbura* 'fly' – *zboară*. In view of the behaviour of *cură* and dialectal *țese/coase*, it now seems preferable to say that the set of diachronic developments observed for this particular group does not rely on analogical pressure from specific verbs but reflects a more general tendency in the language. However, while the situation which resulted in the above cases can be identified synchronically as heteroclisis, it is not clear that the same is true of the *omori* type. This is because, while it may be appropriate to treat some forms in the paradigm as bearing first conjugation morphology, it is difficult to justify the identification of the remainder with any other existing class. As stated above, these verbs originate in the fourth conjugation, and some presentations continue to treat them as comprising a minor subclass of this conjugation (e.g. Lombard 1974, Avram 2001): this makes sense for historical reasons but also for purposes of exposition, as the inflectional behaviour of the augmented *-î* verbs, at least, can be characterized perfectly as involving normal fourth conjugation morphology but with the replacement of the vowels *i*, *e*, *ea* by their centralized counterparts *î*, *ă*, *a* throughout the paradigm. However, in the absence of the earlier 'centralizing' /r/ which produced the change, there is no conditioning environment to motivate allomorphy of this kind within a single

¹¹⁹ This also ultimately reflects a loanword from Slavic, namely Serbian *pogorje* 'mountainous terrain' (Schulte 2005: 380); note that an earlier variant *a pogori* also continues to exist.

¹²⁰ Besides *a omori*, only *a cobori/pogori* 'go down' together with derived *a scobori* 'id.', *a vâri* 'thrust', *a tăbări* 'to inveigh (against someone)' and *a obori/dobori* 'fell' behave in this way.

conjugation: although all verbs in *-î* do have stems in *r*, regular fourth conjugation verbs in *-ri* are numerous, e.g. *a muri* ‘die’ (unaugmented), *a dori* ‘desire’ (augmented). In fact, by the same logic it is equally appropriate in synchronic terms to take the unaugmented verbs in *-î* as variant *first* conjugation forms: the paradigm of unaugmented *a omorî* now patterns perfectly with that of *a cânta*, *modulo* the identity of the theme vowel where this appears, as can be seen by comparing the paradigm of *a omorî* below with that of *a cânta* seen in Table 4.1 at the outset of this chapter (cf. Juilland and Edwards 1971: 104). This is true precisely as a result of the conjugational developments in the N-pattern forms just discussed, which distance the inflectional behaviour of this verb from that of the fourth conjugation.

Table 4.18 The unaugmented ‘fifth conjugation’ in *î*

Infinitive *omorî*
Gerund *omorând*
Past participle *omorât*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>omór</i>		<i>omór</i>	<i>omorám</i>	<i>omorái</i>	<i>omorásem</i>
2sg	<i>omóri</i>	<i>omoárá</i>	<i>omóri</i>	<i>omorái</i>	<i>omoráși</i>	<i>omoráseși</i>
3sg	<i>omoárá</i>		<i>omoáre</i>	<i>omorá</i>	<i>omorí</i>	<i>omoráse</i>
1pl	<i>omorám</i>		<i>omorám</i>	<i>omorám</i>	<i>omorárám</i>	<i>omoráserám</i>
2pl	<i>omoráți</i>	<i>omoráți</i>	<i>omoráți</i>	<i>omoráți</i>	<i>omoráráți</i>	<i>omoráseráți</i>
3pl	<i>omoárá</i>		<i>omoáre</i>	<i>omoráu</i>	<i>omoráră</i>	<i>omoráseră</i>

In fact, as elsewhere in Romanian the identity of the theme vowel is the chief feature distinctively marking membership of a particular conjugation class, it may seem most appropriate not to associate the verbs in *î* with either of these well-established conjugations, but to treat them as constituting a fifth conjugation with its own specific theme vowel, despite the small number of lexemes involved: this is the approach taken by, among others, Juilland and Edwards 1971 as well as Nedelcu (2013: 22f.) in the most recent version of the Romanian Academy Grammar. But this entails that the label of ‘heteroclisis’ cannot strictly be applied to this case. If 1pl and 2pl present *omorám*, *omoráți* can be taken as displaying regular fifth conjugation

morphology, then there is no reason not to say the same for 3pl present *omoarǎ*, whose form therefore does not contain first conjugation morphology out of place: we are simply in the presence of the regular set of fifth conjugation forms.

This case thus provides an illustration of the logical independence of conjugational incursion, as a diachronic process, from heterocclisis as a label for a synchronic state of affairs: while incursion by its nature might be thought to lead ultimately to heterocclisis, we see here one way in which the outcome of incursion may not be best characterized in these terms. Indeed, as previously mentioned, the adoption of innovative endings in the N-pattern cells is one factor making it less appealing to continue to associate the verbs in *-î* with the fourth conjugation synchronically: if every form of *a omorî*, like every form of *a urî*, had continued to pattern with the corresponding form of regular fourth conjugation verbs in *-i*, it would be hard to avoid subordinating this small group of verbs to the fourth conjugation despite the lack of a conditioning environment licensing the centralized vowels in the modern language.

The point of view which assigns the verbs in *-î* to a conjugation of their own, and thus rules out heterocclisis here, is not the only one possible. For one thing, given the small size of the group of verbs involved – and particularly of the group of unaugmented verbs in *-î* – it is hard to imagine that the association of their conjugational morphology with that of the large and productive first and fourth conjugations, which is diachronically unmistakable, goes synchronically unrecognized by speakers. It can also be noted that, as a lexical matter, close links remain between the verbs in *-î* and the fourth conjugation in *-i*, as several items of

vocabulary¹²¹ may appear in either guise, displaying ongoing social variation; this continues an earlier situation where such variation was more frequent (Diaconescu 1977: 73).

Nonetheless, we can say that the behaviour of verbs such as *a omorî* in the modern Romanian standard is not fully satisfying as an example of heteroclisis: it could be considered as highly non-canonical in terms of the framework laid down in 2.3, along a dimension involving the extent to which the inflectional classes ostensibly providing the material for the heteroclitite group can be said to exist independently.

There are fewer doubts with regard to the final phenomenon to be discussed in this chapter, which unambiguously implicates both the first and the fourth conjugations whether viewed in terms of diachrony or synchrony. For an example of the conjugational behaviour in question, in Table 4.19 I provide the standard Romanian paradigm of *a acoperi* ‘cover’:

Table 4.19 The paradigm of *a acoperi* ‘cover’ in standard Romanian

Infinitive *acoperi*
Gerund *acoperind*
Past participle *acoperit*

	present	imperative	subjunctive	imperfect	preterite	pluperfect
1sg	<i>acópăr</i>		<i>acópăr</i>	<i>acopereám</i>	<i>acoperi</i>	<i>acoperisem</i>
2sg	<i>acóperi</i>	<i>acóperă</i>	<i>acóperi</i>	<i>acopereái</i>	<i>acoperişi</i>	<i>acoperiseşi</i>
3sg	<i>acóperă</i>		<i>acópere</i>	<i>acopereá</i>	<i>acoperi</i>	<i>acoperise</i>
1pl	<i>acoperím</i>		<i>acoperím</i>	<i>acopereám</i>	<i>acoperirăm</i>	<i>acoperiserăm</i>
2pl	<i>acoperiţi</i>	<i>acoperiţi</i>	<i>acoperiţi</i>	<i>acopereáţi</i>	<i>acoperirăţi</i>	<i>acoperiserăţi</i>
3pl	<i>acóperă</i>		<i>acópere</i>	<i>acopereáu</i>	<i>acoperiră</i>	<i>acoperiseră</i>

The distribution of conjugational behaviour over this paradigm is by now familiar: while for the most part *a acoperi* patterns regularly according to the fourth conjugation, *acóperă* in the 3sg present, 3pl present and 2sg imperative and *acópere*

¹²¹ At least *a ocari/ocari* ‘scold’, *a oțari/oțari/oțeri/oțeri* ‘snap at’, *a posomori/posomori* ‘sadden’, *a târi/târâi* ‘drag’, *a zădări/zădări* ‘provoke’, *a zăvori/zăvori* ‘bolt (tr.)’, cited by Schulte (2005: 387) from Coteanu et al. 1998.

in the 3sg/3pl subjunctive are unexpected – but would be regular for the given cells in a first conjugation verb. Naturally, this means that they also resemble the corresponding forms in *a omorî*. Lombard wonders (1955: 753f.), therefore, whether 3sg present *acóperă*, as well as analogous *(des)cóperă* and *súferă* from *a (des)coperi* ‘(un)cover’ and *a suferi* ‘suffer’, might share an origin with *omoáră* in the centralizing effect of the preceding consonant. These verbs continue Latin **(ad-/dis-)cooperīre* and **sufferīre*, and thus did not inherit /rr/; however, this does not entirely rule out centralization, which occasionally arose even where it might not be expected on the basis of the Latin etymon (cf. *a amarî* ‘make bitter’, derived from *amārus* ‘bitter’), and Lombard concludes that it is impossible to say whether the -ă of these forms originates in sound change.

However, the obvious objection arises that the rest of the paradigm of these verbs shows no sign of centralization: we do not see infinitive **acoperî* or imperfect **acoperam*. What is more, in early attestations these verbs behave entirely unlike *a omorî*. While the latter has already adopted first conjugation morphology in the 3pl present *omoară* by the sixteenth century, Zamfir’s corresponding records for *a acoperi* (2005: 206-10) show no sign of any irregularity in the present even as late as 1700: the expected fourth conjugation forms 3sg *acopere*, 3pl *acopăr(u)* are universally found. Accordingly, there is no reason to think that centralization in the 3sg present has played any role in the development of heteroclite inflection in verbs of this type.

In fact, the most plausible account reverses the ‘direction’ both of morphological influence within the paradigm and of the sound change involved. Synchronically unusual inflectional behaviour in *a acoperi* can first be identified in the 3sg/3pl *subjunctive*, whose modern form *acopere*, unexpected for a fourth

conjugation subjunctive because of its final *-e*, is already the norm at the time of these early attestations (similarly *descoapere*, *sufere*). This has been explained as the reflex of regular **(ad-/dis-)cooperiat*, **sufferiat* by way of a form with a palatal *r*, which *fronted* the vowel of the ending, cf. *căldare* ‘bucket’ < Late Latin *caldāria* ‘cooking pot’.¹²² It is this subjunctive form in *-e* which appears to trigger the incursion of first conjugation morphology into the paradigm of verbs of this type, just as the 3sg present in *-ă* triggers it in *a omorî* and the other examples seen above. Thus the fronting of a central vowel is crucial to the beginnings of this morphological shift, just as the centralization of a front vowel has been observed to be crucial to those treated above.

This development is interesting, because it provides further support for the notion that the cells involved operate as a morphologically coherent zone within the paradigm: in this case, aberrant behaviour in the third person *subjunctive* cells has been sufficient to give rise to morphological innovation in the third person *present* cells, although one would not expect the subjunctive to exert analogical pressure on the present on the grounds of salience or frequency of usage. The fact that the analogical shift took this counterintuitive direction suggests strongly that speakers do not ‘object’ to the presence of heteroclisia in the verb paradigm; rather, in preserving the alternation between fourth and first conjugation morphology in this case, speakers have redistributed it in accordance with a pattern already observed to govern paradigmatic alternations. The pre-existing example of the *omorî* pattern is also likely to have provided support for the distribution of first and fourth conjugation morphology in this way: if this is the case, it entails that the covert

¹²² Zamfir (2005: 207) follows this explanation, which can be traced back to Tikin (1905: §237.2).

fourth conjugation identity of the ‘fifth conjugation’ is indeed still recognized by speakers.

More generally, though, it must be recognized as noteworthy that although the development seen here in *a acoperi* has in one sense gone in a different ‘direction’ through the paradigm from those developments seen previously which have come about thanks to the morphologization of sound change, in another sense all of these phenomena share in the same direction of travel through the conjugational system – in that they have all moved towards the adoption of first conjugation morphology where none was present before. It is not clear why this should have been the case. In light of the fact that the Romance languages generally continue the especial productivity and populousness of the first conjugation seen in Latin (cf. Aronoff 1994: 43-47 and section 3.3 above), it may seem reasonable to treat the trajectory observed in these several Romanian instances as reflecting a move towards the unmarked or ‘natural’ situation, along the lines of the diachronic move towards weak verbal inflection in the German system treated by Bittner 1985, 1996 and discussed in section 2.2.2. However, as was noted in the introduction to this chapter, while the first conjugation is indeed the most productive in modern Romanian, this was not the case for much of its history: the *fourth* conjugation prevailed (and was, for example, the chief recipient of loan vocabulary) up to the nineteenth century, meaning that the rise of the first conjugation came too late to account for the distinctive pattern of conjugational change first attested here in *a omorî* and the like – not to mention that these verbs reliably fail to show the productive first conjugation inflectional augment, even though they move to the first conjugation in precisely those N-pattern cells which would display it.

Instead, Maiden (2009: 73) raises the possibility that what is going on in these cases, in the terms of Carstairs-McCarthy 1994, is a question of the avoidance of *blur* in affixal function. To take the example of the 3sg present endings: before the application of the centralizing sound change which produces forms such as dialectal *coasǎ*, we can say that *-ǎ* is a class-identifier of the first conjugation while *-e* is the default applied to other conjugations; but once *coasǎ* is found, *-ǎ* appears to be available to mark 3sg present in the third conjugation as well, constituting a situation of inflectional blur in which neither ending identifies a specific class. Accordingly, the shift of conjugational affiliation in *coasǎ* could be seen as the inflectional system's response to this intolerable situation: if the form *coasǎ* itself can be treated by the inflectional system as belonging exceptionally to the first conjugation, then the ending *-ǎ* retains its association with the first conjugation in particular and No Blur is satisfied. However, it seems clear that if individual words are licensed, within the No Blur approach, to alter their inflectional class in particular cells when their forms prove impossible for the Principle to capture, then it is rendered void of content. Any form not satisfying No Blur could be analysed as underlyingly behaving in this way, even in the absence of any independent motivation for positing inflectional change in the given form (as is provided here by the subsequent changes observed elsewhere in the paradigm).

The pattern shown by *a acoperi* is also notable for theoretical reasons, because it seems to provide an example of a type of heteroclisis which cannot be handled in terms of the framework provided by Noyer 2005 for dealing with 'systematic mixed inflection' by means of interclass syncretism, as discussed in section 2.4. I recap the essentials here. Systematic mixed inflection is illustrated in Table 4.20 below as characterizing a 'special' subclass of inflectional class I: this

subclass unexpectedly shares the piece of inflectional material *y* with at least one other inflectional class, in opposition to the majority of class I which makes use of the inflection *c* to realize the same property (case 3 in this schema). Noyer proposes that, where systematic mixed inflection occurs, and cannot be explained in terms of phonological conditioning, the inflectional material symbolized as *y* can only be one of two things, owing to the internal logic of the model. Either it is the *default inflection* for case 3; or the special subclass of class I is characterized by a feature [+F], and *y* is the specific realization of case 3 for the *default class* joined by lexemes bearing the feature [+F].

Table 4.20 Schema for systematic mixed inflection (cf. Tables 2.12 and 2.13)

	class I	class I 'special'	class II	class III
case 1	<i>a</i>	<i>a</i>	<i>w</i>	<i>w</i>
case 2	<i>b</i>	<i>b</i>	<i>x</i>	<i>v</i>
case 3	<i>c</i>	<i>y</i>	<i>y</i>	<i>y/z</i>

In Table 4.21, the relationship between the heteroclite *acoperi* type and the regular conjugations of Romanian is laid out schematically in accordance with Noyer's presentation. If the *acoperi* type is to be treated as a special subclass of any conjugation, this must be the fourth, as the large majority of its forms pattern with this conjugation: the few instances of first conjugation morphology can be treated as intrusive both synchronically and diachronically. The infinitive and 2pl present endings are given in the table as examples of the very many inflections in which *a acoperi* and the fourth conjugation pattern together and in opposition to the other conjugations (cf. case 1 and case 2 in Noyer's schema). Meanwhile, the 3sg present endings are given to stand for the few parts of the paradigm in which the fourth conjugation inflection is not shared by the 'special' *acoperi* type (cf. case 3 in Noyer's schema).

Table 4.21 The Romanian conjugations and *a acoperi*

conjugation	4	4 'special'	1	2	3
	<i>dormi</i>	<i>acoperi</i>	<i>cânta</i>	<i>putea</i>	<i>ţese</i>
infinitive	-i	-i	-a	-ea	-e
2pl present	-iţi	-iţi	-aţi	-eţi	-eţi
3sg present	-e	-ă	-ă	-e	-e

The point is this. As *a acoperi* disagrees with the regular fourth conjugation by taking the ending *-ă* in the 3sg present, which is instead shared with the first conjugation, the Distributed Morphology (DM) model of systematic mixed inflection allows for two possibilities. The first is that *-ă* is the default inflection for the 3sg present in Romanian, and it is the *other*, apparently 'non-special', fourth conjugation verbs whose ending *-e* is specified. However, this solution is clearly untenable, as the ending *-e* is not specific to the normal fourth conjugation: it is also shared by the second and third conjugations, and must represent the default inflection here in DM terms. The second is that in the 3sg present *a acoperi* has become impoverished of the inflectional class feature which assigns it to the fourth conjugation, but that it bears some *morphosyntactic* feature which means that in the absence of an inflectional class feature it is reassigned to the first conjugation. But no such morphosyntactic feature, analogous to gender in the nominal examples treated by Noyer, can be posited for verbs. There is thus no way to account for the heteroclisis shown by *a acoperi* as an example of systematic mixed inflection.

In DM terms, then, the inflection of *a acoperi* must be considered idiosyncratic – 'heteroclisis' in the restricted sense. This suggests that it could never become adopted by speakers as a general model of conjugational behaviour to which new lexemes could become attracted: the resulting subclass of the fourth conjugation could not be generated by the linguistic system. However, the evidence shows that the pattern has not in fact remained restricted to the few verbs in which it originally

arose as discussed above. Romanian possesses several loanwords from Italian or French ultimately reflecting Latin prefixed verbs in *-ferō*: of these, *a oferi* has a fourth conjugation infinitive thanks to Italian *offerire*, but *a deferi* and *a diferi* take these forms purely on the model of *a suferi*, with which they have evidently been taken to share a root. All three of these verbs also share the heterocclisis shown by *a suferi*. Moreover, Lombard reports (1955: 620, 626-28) that *prefer-*,¹²³ *refer-*, and *confer-*, for which full first conjugation and fourth conjugation paradigms can in principle be found (reflecting the pressure of *a oferi* etc. and the fact that they are borrowed from French *préférer* etc.), tend to assign morphology from the two conjugations in the same distribution over the paradigm, suggesting it represents a compromise position between the two; other neologisms based on Latin roots may behave likewise, e.g. *a absolvi/-a*. Verbs with a longer tenure in Romanian may also show the same pattern, e.g. *a curăți/-a* ‘clean’ (from native *curat* ‘clean’), *a gădili/-a* ‘tickle’,¹²⁴ and notably *a sprijini* ‘support’ (ultimately from Slavic), which has never had a full first conjugation paradigm; it may be relevant that these verbs, whose behaviour cannot be explained as the result of general tendencies affecting modern loanwords, share the accentual pattern of *a acoperi*, stressing the penultimate syllable of the root (thus 1sg present *cúrăț*, *gădil*, *sprijin*, cf. *acópăr*).¹²⁵

It appears, then, that heterocclisis bringing together both fourth and first conjugation morphology has not only been retained in those verbs in standard Romanian where it originally arose as the result of the morphological reinterpretation of phonologically regular forms; the distribution of these two types of inflectional behaviour over the paradigm has been generalized to lexemes to which it did not

¹²³ Regular *a prefera*, with no fourth conjugation morphology, is now the norm for this verb.

¹²⁴ This verb is now normally found with first conjugation forms throughout.

¹²⁵ Presumably with such forms in mind, Feldstein 2004 aims to reduce the unexpected conjugational behaviour of all verbs of this type to a purely phonological effect of stress placement. However, this cannot account for the behaviour of *a oferi*, *a absolvi* etc. (1sg *ofér*, *absólv*).

originally apply, or even which did not previously exist in the language, for a wide variety of reasons. This observation provides robust evidence against a model of inflectional morphology which says that the pattern in question should be not only disfavoured but impossible for the linguistic system to generate as a principled type of morphological behaviour.

4.5 Conclusion

The investigation presented above, treating the origins and synchronic patterning of heteroclisis and associated inflectional behaviour in Daco-Romanian varieties, has naturally involved both evidence and argumentation somewhat different from that provided for the Latin cases treated in the previous chapter, where internal and comparative Indo-European reconstruction played a central role alongside textual attestations. In examining the history of the Romanian verb system, we are able to benefit both from our understanding of its historical background in Latin itself and from the copious available (although in part unpublished) documentation of inflection in its dialects, down to the level of individual verb forms as found in individual villages.

This wealth of data allows us to observe instances of heteroclisis appearing at the dialectal level in verbs which inflect regularly, in accordance with their etymologically expected conjugation, in the standard language: (the local equivalents of) *a țese* ‘weave’ and *a coase* ‘sew’ in Oltenia, *a sughița* ‘hiccup’ in Banat, and *a zbură* ‘fly’ in Crișana can all be mentioned here. However, heteroclisis is by no means a dialectal feature absent from the standard; on the contrary, the heteroclitic behaviour exemplified here by standard *a acoperi* ‘cover’, originally restricted to just a few verbs as the result of a morphological reinterpretation of a particular sound change, has apparently come to provide a model for the redistribution of conjugational morphology over the paradigm of numerous neologisms and inherited forms to which this sound change never applied – a development which seems to counterexemplify Noyer’s (2005) position on ‘systematic mixed inflection’.

As seen previously for Latin, one way for heteroclisis to develop is as a result of neutralizing sound change, which brings together previously distinct patterns of

inflectional marking in a portion of the paradigm of certain verbs. This effect is again visible in several of the developments investigated in the present chapter. However, in these Romanian cases we are in a better position to see evidence that the products of these phonological developments have been reinterpreted as morphologically innovative: it is perhaps clearest from the different patterns of heteroclisis visible over the paradigm of *a țese/a coase* in different local varieties that, thanks to a sound change (vowel centralization) inherently ‘blind’ to morphological structure, speakers have identified first conjugation morphology where it was not originally found, and have subsequently extended it to parts of the paradigm where it cannot have arisen as the direct result of the initial sound change. Meanwhile, again as was suggested for Latin, we see evidence from Romanian of the potential for heteroclisis to emerge as the result of the influence of individual lexical items on one another, whether on the basis of a semantic relationship (*a zbură* influenced by *a sări*), a formal relationship (*a mâna*, *a amâna* influenced by *a rămănea*), or perhaps both at once (*a sughița* influenced by *a înghiți*). The redistribution of rival productive forms along the lines laid down by pre-existing heteroclite lexemes, as I suggest for neologisms such as *a referi*, *a absolvi*, represents a further origin for heteroclisis.

The details of the developments treated in this chapter also demonstrate the part that can be played by purely morphological or ‘morphomic’ patterns, to which no coherent morphosyntactic motivation can be attributed, in the articulation of the inflectional paradigm. The emergence of ‘L-pattern’ alternations in verbs where they are not expected on historical grounds (*a zbură*, *a (a)mâna*) can be explained by the direct morphological influence of specific verbs where this pattern of stem-final consonant alternation is already observed; it is rather more striking that we can observe the adoption of the ‘N-pattern’ distribution in newly heteroclite verbs from

various sources in both the dialects and the standard, pointing to its underlying importance to the structure of the verb paradigm. The robustness of this pattern as a productive model for innovation is demonstrated by the fact that in *a acoperi* it seems to have led to conjugational change in present (indicative) forms on the basis of a reinterpretation of the third person subjunctive, a directionality which would not be expected on functional grounds. Meanwhile, the existence of a ‘PYTA’ distribution for heteroclisis in *a țese/a coase* in central and southern Oltenia, which must represent a further development on the basis of the N-pattern distribution visible in the north of the region, illustrates the potential for close dialectological study to uncover the diachrony of morphological change.

Finally, in the history of the ‘fifth conjugation’ represented here by *omorî* ‘kill’, we have also seen a case which shows that the diachronic incursion of innovative conjugational behaviour into the paradigm need not necessarily give rise to heteroclisis in the resulting synchronic system: this bears out the principle, as stated in previous chapters, that although one is often responsible for the other, the two phenomena are logically independent. In the following chapter, which treats heteroclisis in Romansh, an ongoing development will be presented from the Surselvan dialect (section 5.3) which suggests a different way in which what appears to be the diachronic counterpart of heteroclisis may not give rise to heteroclisis in synchrony, but may instead lead to a new model of inflection which cannot be taken as ‘parasitic’ on other inflectional classes.

5. Heterocclisis in Romansh

5.1 Introduction

This chapter treats the phenomenon of heterocclisis as it is found in the verb systems of the group of Romance varieties which together constitute Romansh.

Spoken in the Swiss canton of Graubünden, Romansh is now the native language of no more than 50,000 people, less than 1% of the population of Switzerland as a whole. Practically all of these are bilingual in German, and indeed the number of Romansh speakers is so small precisely because for many centuries inhabitants of the region have increasingly turned to German as the local prestige language, leaving only the least urbanized areas as predominantly Romansh-speaking territory; the last urban centre to become Germanized was Chur, in the fifteenth century (Haiman and Benincà 1992: 11), and in the present day Romansh is no longer spoken over a contiguous area but is interrupted by monolingual German zones.

This long-standing sociolinguistic trend has led to a state of affairs in which the local varieties descended from Latin do not share in a common affiliation with a single standard language emanating from a centre of political and cultural influence.¹²⁶ Instead, five separate dialect areas have emerged, each with its own local standard: from west to east, these are Surselvan, Sutselvan, Surmeiran, Puter and Vallader. The last two are also known as Upper and Lower Engadine respectively and are commonly referred to together as the Engadine dialects; the remainder are the Rhenish dialects, of which Surselvan (with around 18,000 speakers) is by far the most prominent, while smaller Sutselvan and Surmeiran are

¹²⁶ The wholly artificial, overarching standard *Rumantsch Grischun* was developed by Heinrich Schmid (1982) as a written language only, to represent Romansh (as Switzerland's fourth official language) in government documents.

sometimes taken together as Central Romansh. Of the many distinctions at all linguistic levels which separate these five dialects, some involve the verb paradigm and will accordingly be touched on here to the extent that they are relevant; in particular, sections 5.2 and 5.3 deal with developments involving heteroclisis which are restricted to particular dialects. However, it should be emphasized that the varieties labelled as ‘Romansh’ do belong together genetically and typologically as well as geographically, and have more in common with each other than with other Romance varieties of the area (see e.g. Prader-Schucany 1970, on isoglosses between Romansh and the Lombard varieties spoken to the south). It is much less clear whether the same can be said for the putative ‘Rhaeto-Romance’ of which Romansh is often said to form a part, together with Dolomitic Ladin and Friulian (after Ascoli 1873, 1882-85): indeed, the authors of the standard manual *The Rhaeto-Romance Languages* themselves acknowledge that the upshot of their survey is to cast doubt on the idea that any such unit exists.¹²⁷ The present study deals only with the more restricted and coherent entity of Romansh.

The relevance of the concept of heteroclisis to the description of this language’s verb morphology has been highlighted only recently, by Maiden (2009: 83f.), in a very brief presentation comparing the most widespread type of heteroclite behaviour seen in Romansh (in which characteristic fourth conjugation morphology has infiltrated the paradigm of some traditionally third conjugation verbs) with the independent Romanian instances dealt with at greater length elsewhere in his article (and discussed in Chapter 4 above). I begin here (section 5.2) by providing a fuller

¹²⁷ ‘There are no very convincing reasons for grouping together as a single language the various dialects known as Rhaeto-Romance.’ (Haiman and Benincà 1992: 8). This wording (‘as a single language’) is somewhat misleading, as the context makes clear that the question at issue is in fact whether Romansh, Ladin and Friulian make up a single genetic clade to the exclusion of all other Romance varieties. The present chapter does not privilege the other ‘Rhaeto-Romance’ varieties with any special status vis-à-vis Romansh.

account of the same phenomenon, whose effects are seen in all Romansh dialects, and point out its relevance for our understanding of the Romansh verb system more generally: as in Romanian, the characteristic patterns taken up by heteroclite inflection reveal the existence of morphological patterns within the paradigm, which are known implicitly to speakers and which channel the spread of morphological change. In order to do this I first provide an overview of Romansh verb inflection.

In the two sections which follow, I move on to discuss two more recent developments which are restricted to particular parts of the Romansh-speaking territory. Section 5.3 draws attention to a striking and apparently ongoing change affecting Surselvan, in which the originally exceptional behaviour addressed in 5.2 has been lexically generalized to the extent that it is now becoming the new inflectional norm: this has repercussions for the description of the Surselvan conjugations and provides a contemporary Romance illustration of the necessity of distinguishing between synchronic and diachronic claims about heteroclitism. Section 5.4, meanwhile, deals with a heteroclite phenomenon unique within Romansh and within the Romance family as a whole, in which verbs of all conjugations have come to show distinctive third conjugation behaviour in the infinitive: I demonstrate that this is a genuinely morphological phenomenon.

5.2 Pan-Romansh heteroclisis: fourth conjugation morphology ‘out of place’

Although Maiden 2009 is, to my knowledge, the first discussion to use the term ‘heteroclisis’ itself with reference to the behaviour of the Romansh verb, the phenomenon treated in this section did not go unnoticed in earlier work on the Romansh dialects. The verbs of Romansh are usually separated into four conjugations, on the basis of distinctions (in the infinitive and elsewhere) which survive from the ancestral four-conjugation system of Latin, but most descriptions, normative grammars and textbooks of the dialects also make some mention of the fact that certain anomalous verbs, which are assigned on the basis of their infinitive to the third or second conjugation, display unexpected fourth conjugation morphology in some parts of their paradigm. However, the marginal nature of this behaviour within the conjugational system as a whole¹²⁸ means that sample paradigms or full accounts detailing the inflection of such verbs are rarely provided. A notable exception is Arquint’s (1964) pedagogical description of Vallader (followed by Maiden 2009), which on pp. 134f. summarizes the inflectional behaviour of this group of verbs. The paradigmatic behaviour presented by Arquint for these heteroclite verbs in Vallader tallies, *mutatis mutandis*, with the generally more cursory remarks made by authorities on other dialects and with the evidence of textual attestations from across Romansh supplied by the monumental *Dicziunari rumantsch grischun* (DRG, von Planta and Melcher 1939-). Moreover, the verb system of Vallader is relatively free of distinctive dialectal features in terms of its morphology or phonology which concern the identity of the regular conjugations (unlike Surselvan and Surmeiran respectively), meaning that the behaviour of the

¹²⁸ In most dialects at least; but see section 5.3 for evidence that in Surselvan the behaviour generally treated as marginal in fact represents the innovative norm for regular verbs of the second and third conjugations.

Vallader heteroclites can be treated as representative of the basic form taken by the Romansh phenomenon as a whole. In this section I thus illustrate this phenomenon with data taken from Arquint, for the Vallader verb *cuorrer* ‘run’ (Table 5.4 below).

However, the notable features of this type of heterocclisis can only be appreciated in the context of the regular behaviour of Romansh verbs. I therefore begin by providing an outline of this behaviour and its historical background, again taking Vallader as a proxy for ‘common Romansh’. Tables 5.1, 5.2 and 5.3 (also after Arquint 1964) give the principal synthetic forms in the paradigm of three regular verbs in Vallader,¹²⁹ belonging to the first, third and fourth conjugations respectively, and exemplify most of the characteristics of Romansh verb inflection which will be relevant here. Stress is marked here with an acute accent, which is not present in the standard orthography; the term ‘alternating’ is explained below. As elsewhere in Romance, the productive conjugations are the fourth and (especially) the first; meanwhile, the second and third largely contain verbs inherited from those conjugations in Latin.

Table 5.1 Paradigm of a regular (‘alternating’) first conjugation verb

inf. *nodár* ‘swim’
 pres.pt. *nodónd*
 past pt. *nodá*

	pres. ind.	imper.	pres. subj.	impf.	impf. subj.	perfect	future
1sg	<i>noud</i>		<i>nóuda</i>	<i>nodáiva</i>	<i>nodéss</i>	<i>nodét</i>	<i>nodará</i>
2sg	<i>nóudast</i>	<i>nóuda</i>	<i>nóudast</i>	<i>nodáivast</i>	<i>nodéssast</i>	<i>nodéttast</i>	<i>nodarást</i>
3sg	<i>nóuda</i>		<i>nóuda</i>	<i>nodáiva</i>	<i>nodéssa</i>	<i>nodét</i>	<i>nodará</i>
1pl	<i>nodáin</i>		<i>nóudan</i>	<i>nodáivan</i>	<i>nodéssan</i>	<i>nodéttan</i>	<i>nodarán</i>
2pl	<i>nodáivat</i>	<i>nodái</i>	<i>nóudat</i>	<i>nodáivat</i>	<i>nodéssat</i>	<i>nodéttat</i>	<i>nodarát</i>
3pl	<i>nóudan</i>		<i>nóudan</i>	<i>nodáivan</i>	<i>nodéssan</i>	<i>nodéttan</i>	<i>nodarán</i>

¹²⁹ Second conjugation verbs are only distinct from third conjugation verbs in the infinitive, which has the stressed ending *-áir*, e.g. *placháir* ‘please’. I continue to refer to second and third conjugation verbs separately, for ease of reference across the thesis as a whole, as the point does not affect my presentation.

Table 5.2 Paradigm of a regular third conjugation verb

inf. *víver* ‘live’
 pres. pt. *vivónd*
 past pt. *vivú*

	pres. ind.	imper.	pres. subj.	impf.	impf. subj.	perfect	future
1sg	<i>viv</i>		<i>víva</i>	<i>viváiva</i>	<i>vivéss</i>	<i>vivét</i>	<i>vivará</i>
2sg	<i>vívast</i>	<i>víva</i>	<i>vívast</i>	<i>viváivast</i>	<i>vivéssast</i>	<i>vivéttast</i>	<i>vivarást</i>
3sg	<i>víva</i>		<i>víva</i>	<i>viváiva</i>	<i>vivéssa</i>	<i>vivét</i>	<i>vivará</i>
1pl	<i>viváin</i>		<i>vívan</i>	<i>viváivan</i>	<i>vivéssan</i>	<i>vivéttan</i>	<i>vivarán</i>
2pl	<i>viváivat</i>	<i>vivái</i>	<i>vívat</i>	<i>viváivat</i>	<i>vivéssat</i>	<i>vivéttat</i>	<i>vivarát</i>
3pl	<i>vívan</i>		<i>vívan</i>	<i>viváivan</i>	<i>vivéssan</i>	<i>vivéttan</i>	<i>vivarán</i>

Table 5.3 Paradigm of a regular (‘augmented’) fourth conjugation verb

inf. *finír* ‘end’
 pres. pt. *finínd*
 past pt. *finí*

	pres. ind.	imper.	pres. subj.	impf.	impf. subj.	perfect	future
1sg	<i>finísch</i>		<i>finíscha</i>	<i>finíva</i>	<i>finíss</i>	<i>finít</i>	<i>finirá</i>
2sg	<i>finíschast</i>	<i>finíscha</i>	<i>finíschast</i>	<i>finívast</i>	<i>finíssast</i>	<i>finíttast</i>	<i>finirást</i>
3sg	<i>finíscha</i>		<i>finíscha</i>	<i>finíva</i>	<i>finíssa</i>	<i>finít</i>	<i>finirá</i>
1pl	<i>finín</i>		<i>finíschan</i>	<i>finívan</i>	<i>finíssan</i>	<i>finíttan</i>	<i>finirán</i>
2pl	<i>finívat</i>	<i>finí</i>	<i>finíschat</i>	<i>finívat</i>	<i>finíssat</i>	<i>finíttat</i>	<i>finirát</i>
3pl	<i>finíschan</i>		<i>finíschan</i>	<i>finívan</i>	<i>finíssan</i>	<i>finíttan</i>	<i>finirán</i>

As is clear from these paradigms, the synthetic forms employed in the Romansh verb system can generally be traced back to their Latin precursors, or else recognized as post-classical formations also represented in other Romance languages. Examples of the former include the infinitive, present participle (from the Latin gerund), imperfect, and imperfect subjunctive (from the Latin pluperfect subjunctive, thus a counterpart to the Romanian pluperfect in terms of its ancestry). Examples of the latter include the future formation, which continues a syntagm of infinitive + inflected forms of ‘to have’ (cf. French *je finirai* and its equivalents in the majority of Romance languages), and the synthetic perfect: retained only in the Engadine dialects, this has been entirely remodelled on the basis of 3sg forms of the second and third conjugations (*vivét* here), which ultimately go back to the late Latin

vendedit type that also underlies Italian *vendetti*, *perdetti* etc. (Stimm and Linder 1989: 775, following Lausberg 1972: 255-58).

Of particular interest here, because of its relevance to the question of conjugation class membership in Romansh, is another salient post-classical feature of the verb system, seen in forms such as 1sg *finisch* [fi'niʃ] – namely the inflectional augment descended from the Latin derivational marker *-ēsc-/-īsc-* (all other Romansh dialects in fact show *-esch-*: Haiman and Benincà 1992: 82f.). As in many Romance languages, including Romanian, this element has become a characteristic of regular inflection in the category continuing the Latin fourth conjugation; the large majority of fourth conjugation verbs are ‘augmented’, showing *-isch-* in the paradigm cells shown in Table 5.3, and this is the productive pattern, while only a few dozen fourth conjugation verbs survive which do not require the augment and whose paradigm thus resembles that of *finir* but with no trace of forms in *-isch-* (e.g. *servir*, 1sg *serv*).

Some first conjugation verbs also show an inflectional augment, although the tendency for verbs to require it is far less pronounced than in the fourth conjugation. However, the first conjugation augment seen in Romansh does not continue a derivational element which already shows an affinity with the first conjugation in Latin (as is the case for Romanian); instead, it is descended from the same element *-ēsc-/-īsc-*, and takes the form *-esch-* throughout the dialects (including in Vallader, where it thus differs from the augment seen in the fourth conjugation). Alongside *nodar* 1sg *noud* ‘swim’, then, there are augmented verbs such as *sperar* 1sg *speresch* ‘hope’. Meanwhile, the inflectional augment is almost never found in the paradigm of verbs belonging to the second or third conjugations (Meul 2013: 194).

But along with the relative familiarity of the morphological material used, these paradigms also illustrate a feature which distances the verb morphology of

Romansh substantially from that of Latin: this is the amount of neutralization which has affected inflectional distinctions found in the latter. On grounds of historical phonology we can identify the Latin first conjugation as the ultimate source of the unstressed inflectional endings now shared by all verbs, regardless of conjugation, in the singular and 3pl present indicative and in the singular imperative (neutralization in the present subjunctive proceeded differently in different Romansh varieties).¹³⁰ Meanwhile, the second and third conjugations have fallen together almost entirely, as in most Romance languages, and in several parts of the paradigm these have undergone neutralization with the first conjugation as well, with either the former (e.g. 1pl present indicative *-ain* < *-ēmus*, imperfect *-aiva* etc. < *-ēbam* etc.) or the latter (e.g. present participle *-ond* < *-andus*) originally providing the morphology they now have in common.

Besides these conjugational distinctions, an important formal distinction found in Latin *within* the paradigm of individual verbs has also been effaced in Romansh. Although we have seen that the Latin pluperfect subjunctive is continued in the category labelled here as the ‘imperfect subjunctive’,¹³¹ this no longer brings with it the requirement of a special second or ‘PYTA’ stem, as the equivalent does in (for example) Romanian: instead it is built on the stem found in the (present) infinitive and generally across the paradigm. The same goes for the heavily remodelled perfect forms seen in Engadine. In the case of regular first and fourth conjugation verbs, the lack of a distinct PYTA stem is not unexpected given the state of affairs already seen in Latin, where the present stem *amā-* clearly underlies the perfect stem *amāv-*. However, in Romansh this neutralization within the paradigm

¹³⁰ Prader-Schucany (1970: 167-70, 176-78, 181 fn. 1). In the Rhenish varieties the present subjunctive retains conjugational distinctions in 1pl and 2pl, like the present indicative.

¹³¹ The same form is often labelled as the ‘conditional’, especially in Surselvan where the synchronic present subjunctive endings can be added to it to give an ‘indirect conditional’: see e.g. Spescha 1989.

has been generalized across the lexicon, as illustrated by forms such as *vivess* above (which does not continue Latin *vīx-*, the *perfectum* stem of *vīvō*). Specialized PYTA stems survive only in a few verbs which are also irregular otherwise, such as *esser* ‘be’, imperfect subjunctive *füss*. This is all the more striking for the fact that past participles with distinctive stems often do retain these, though they may be close in form to perfect stems which are lost (e.g. *scrit* < *script-um*, while *scrivess* has been remodelled and has given up the stem seen in *scrips-issem*). Thus the morphological split in the paradigm between present and PYTA forms, which in Chapter 4 was identified as representing a barrier to conjugational change in Romanian, has been abandoned in Romansh.

At the same time, however, the verb system in Romansh, as in other Romance languages, does display the innovative type of stem alternation unknown to Latin which I refer to here and in preceding chapters as the ‘N-pattern’ (cf. Maiden 2009: 83f.). This is exemplified here by the opposition between *nod-* and *noud-* in the paradigm of *nodar* ‘swim’, and the formally more substantial, but identically distributed opposition between *fin-* and *finisch-* in the paradigm of *finir* ‘end’. As discussed in section 2.5 (and references given there), this alternation began life as the result of differential sound change acting on stressed and unstressed versions of what was originally a single (present) stem, in a system where stress was assigned on the basis of phonological principles independent of the verb paradigm as such. Once stress position over the paradigm became synchronically arbitrary upon the loss of Latin vowel length, the assignment of stress to the stem or to the ending (and the corresponding choice of stem form itself) could be taken as a morphological characteristic of verb inflection, and accordingly the same ‘N-pattern’ distribution over the paradigm became available not only to stem pairings which arose thanks to

stress alternation (e.g. the ancestors of *nod-* vs. *noud-*), but also to those which did not originally belong to the same lexeme at all (e.g. the ancestors of *fin-* vs. *finisch-*).

As illustrated in the tables above, these alternations remain associated with stress position up to the present day. In fact, for this reason it is a matter of current debate whether the distinction between N-pattern and non-N-pattern stems should be seen as a genuinely morphological phenomenon at all within the synchronic system of Romansh, or whether phonology alone is responsible for the choice of stem. Treating the verb paradigm of Surmeiran, Anderson shows (2008, 2011) that the assignment of stress to Surmeiran words is entirely predictable given their segmental content, and as stress in turn determines the shape of the stem in alternating and augmented verbs, he therefore suggests that this type of stem alternation effectively takes place outside the purview of the morphological system. Anderson 2013 extends this analysis to other Romansh dialects – although without attempting to demonstrate that stress assignment elsewhere in Romansh can be reduced to segmental content as it can be in Surmeiran. In response, Maiden defends the morphological status of the alternation (Maiden 2008, 2011b): among other arguments he points out that even within Surmeiran the proposed phonological conditioning does not in fact apply across the language as a whole, but only within a morphologically defined sphere, namely the inflectional paradigm of the verb. I shall claim below that the phenomenon of pan-Romansh heteroclitism treated in the current section has a part to play in this debate.

Against the background of these general features of the Romansh conjugational system, Table 5.4 presents a paradigm exemplifying the basic type of heteroclitism found across Romansh: specifically, the paradigm of *cuorrer* ‘run’ as it appears in Vallader (Arquint 1964: 134-36).

Table 5.4 The paradigm of the heteroclite verb *cuorrer* ‘run’ in Vallader

infin. *cuorrer* ‘run’
 pres. pt. *currind* / *currond*
 past pt. *curri* / *cuors*

	pres. ind.	imper.	pres. sbj.	impf	impf. sbj.	perfect	future
1sg	<i>cuor</i>		<i>cuorra</i>	<i>curriva</i>	<i>curriss</i>	<i>currít</i>	<i>currará</i>
2sg	<i>cuórrast</i>	<i>cuorra</i>	<i>cuórrast</i>	<i>currívast</i>	<i>curríssast</i>	<i>curríttast</i>	<i>currarást</i>
3pl	<i>cuorra</i>		<i>cuorra</i>	<i>curriva</i>	<i>curríssa</i>	<i>currít</i>	<i>currará</i>
1pl	<i>currín</i>		<i>cuórran</i>	<i>currívan</i>	<i>curríssan</i>	<i>curríttan</i>	<i>currarán</i>
2pl	<i>currívat</i>	<i>currí</i>	<i>cuórrat</i>	<i>currívat</i>	<i>curríssat</i>	<i>curríttat</i>	<i>currarát</i>
3pl	<i>cuórran</i>		<i>cuórran</i>	<i>currívan</i>	<i>curríssan</i>	<i>curríttan</i>	<i>currarán</i>

As can be seen from a comparison of this table with those preceding, the paradigm of *cuorrer* brings together morphological material which in the regular course of events belongs to opposing conjugations: the infinitive itself unambiguously belongs to the third conjugation, while all the underlined forms pattern just as unmistakably with the fourth.

It is these fourth conjugation forms that are unexpected here, as *cuorrer* is an item of native vocabulary inherited straightforwardly from the third conjugation Latin *currere*. The same is true for the handful of other native Engadine verbs displaying this inflectional behaviour, which include *tscherner* ‘select’ < *cernere*, *crescher* ‘grow’ < *crēscere* and *nascher* ‘be born’ < *nāscere*, excepting the lone *cumparair* ‘appear’ < *comparēre*, which shows unexpected fourth conjugation morphology accompanying an infinitive of the second conjugation. What we see here, then, results from the incursion of fourth conjugation morphology into the paradigm of the third (/second) conjugation: a synchronically heteroclite paradigm has come about thanks to what might be called diachronic heteroclis, in which conjugation change has taken place in a subset of the paradigm of a few exceptional verbs.

This heteroclite behaviour has long been recognized for a small group of inherited verbs in every dialect of Romansh,¹³² and is already attested in our earliest written sources. However, the verbs involved have not historically been identical in every dialect and have not remained constant over time. Reflexes of *currere* in the modern dialects unanimously show the same type of heteroclisism as seen above for Vallader, with a third conjugation infinitive but unexpected fourth conjugation morphology elsewhere: for example, the Surmeiran infinitive *correr* is accompanied by 1pl present *currign* (as opposed to the expected third conjugation **curragn*), and Surselvan *cuorer* has 1pl present *currin* (not **currein*). However, the evidence of the earliest Old Surselvan version of the Bible, dating from 1648, shows that at that period *cuorer* could still operate as a regular third conjugation verb, with e.g. 2sg imperfect *currevas* (not **currivas*; Mourin and de Poerck 1964: 100).¹³³ In contrast, the Old Surselvan text already shows heteroclite behaviour in *scriver* ‘write’ < *scribere* and indeed *viver* < *vivere*, which was given above as an example of a regular verb in Vallader: these do not show fourth conjugation behaviour in Old Engadine (Mourin and de Poerck 1964: 90; 234), and have never acquired it since. Meanwhile, modern reflexes of *cernere* are heteroclite in Engadine and Surselvan, but not in the Central dialects of Sutselvan and Surmeiran which lie between them.¹³⁴

The fact that this pan-Romansh heteroclisism became established before our earliest substantial records of the language begin, together with the disagreement shown by the dialects as to which inherited verbs are implicated in it, makes it

¹³² However, in section 5.3 I give evidence that modern descriptions which continue to recognize only a small group of verbs showing heteroclisism in Surselvan present an overly conservative view.

¹³³ At a later stage in Surselvan, the *i* of the fourth conjugation imperfect and imperfect subjunctive was opened to close *e*, which nonetheless seems to have remained distinct from the open *e* of the second/third conjugation forms (Stimm and Linder 1989: 774, following Huonder 1901: 61f.). However, the identical spelling of the two vowels means that it is no longer possible to identify heteroclisism in these parts of the paradigm on the basis of written forms: see fn. 137 below.

¹³⁴ Standard forms are Engadine *tscherner* *-in(s)*, Surselvan *tscharner* *-in*, but Sutselvan *tschearner* *tscharnagn*, Surmeiran *tscherner* *tscharnagn*.

difficult to identify a starting point or a historical rationale for the phenomenon. As the difference between the expected third conjugation and attested fourth conjugation forms generally involves only the quality of the thematic vowel, it is tempting to look for a phonological account, whereby a regular sound change in a certain position was reinterpreted as a morphological effect relevant to conjugation class membership: as we have seen, just such a sequence of events gave rise to widespread heteroclisis in the paradigm of *a coase* and *a țese* in Romanian dialects. Maiden (2009: 83, fn. 38) seems to assume some such phonological trigger when he suggests that inflectional ambiguity (or ‘blur’ in the terms of Carstairs-McCarthy 1998, which presupposes it) in some cell or cells was responsible for the development of heteroclisis in such verbs. But no sound change is attested which could have affected verbs as different in stem shape as *cuorrer*, *scriver* and *tscherner*; and unlike in Romanian it is hard to identify a phonologically coherent group of heteroclite verbs whose innovative morphological behaviour could later have served as a model to be taken up by others, or a sound change which could have given the requisite ambiguous outcome in any particular cell.

A more likely scenario is that the fourth conjugation morphology at issue did not originally arise accidentally in Romansh as the result of a sound change specific to the language but, at least in a small number of the relevant verbs, was available to speakers on independent grounds having little to do with the nature of Romansh conjugation in particular. The Romance languages show a general tendency to extend fourth conjugation morphology spontaneously and idiosyncratically (i.e. quite apart from any systematic changes in the relationship between conjugation classes and the lexicon) to existing verbs which previously belonged to the second and third, but not vice versa – an asymmetry which is already visible in Latin, where the second and

third conjugations are practically closed classes.¹³⁵ It is probable that the appearance of fourth conjugation variants in verbs such as *tscherner* and *cuorrer* is due simply to this tendency, especially as other Romance varieties show similar developments illustrating the effect of this diachronic trend on the same verbs, cf. French *courir* (Pope 1956: 337). The fact that *cuorrer* optionally shows a distinctive fourth conjugation past participle (*curri* as if < **currītus*), which is not generally true of the heteroclite verbs treated here, further suggests that it originally acquired fourth conjugation morphology not according to any Romansh-internal principle but as a lexical idiosyncrasy.

This proposal, namely that heteroclisism arose in Romansh as a response to a situation of morphological rather than phonological uncertainty, is corroborated by one striking feature of the phenomenon which is rarely dwelt upon in the literature. Across Romansh, the lexical incidence of this type of heteroclisism is highest not among verbs inherited from the Latin third conjugation (the native vocabulary), but among verbs borrowed at a later stage from the third conjugation in Latin or the equivalent in other languages: examples include (from Italian) *divider* ‘divide’, *insister* ‘insist’, and (from Latin) *ceder* ‘yield’, *exprimer* ‘express’, whose phonology shows that they cannot be directly inherited from the parent language.¹³⁶ For such verbs, in fact, heteroclite behaviour seems to be the norm. The only exceptions are transparently complex (prefixed) loans which can be identified with a native simplex third conjugation verb already found in the language, such as borrowed *committer* ‘commit’ alongside inherited *metter* ‘place’: these take on the morphological behaviour shown by the existing stem. It is notable that no corresponding tendency

¹³⁵ With the exception of the productive third conjugation formation in *-scere* (Leumann 1977: 552), which ultimately becomes associated with the fourth conjugation itself, cf. section 4.1 above. On the transfer of verbs into the fourth conjugation in Romance, see e.g. Penny (2002: 173) for Spanish.

¹³⁶ See e.g. Arquint (1964: 134f.), Ganzoni (1977: 131f.); cf. *DRG* for etymologies.

exists for verbs borrowed from the fourth conjugation in Latin (or its descendants) to acquire third conjugation morphology in the course of their transfer into Romansh. Here too, then, we see evidence of a bias towards the adoption of morphology proper to the fourth conjugation at the stage when such learned vocabulary was borrowed into the language. This suggests that heteroclisis arose as the result of a general pressure away from unproductive third/second conjugation morphology in favour of the fourth, but that fourth conjugation behaviour became more easily fixed in certain parts of the paradigm: this echoes what was seen in the previous chapter (end of section 4.4) for modern loanwords into Romanian, which tend to make use of both first and fourth conjugation morphology within the paradigm.

The distribution of forms belonging to the different conjugations over the paradigm of *cuorrer* is itself worthy of comment. Heteroclisis is certainly found, as can be seen simply from the juxtaposition of the third conjugation infinitive *cuorrer* and the 1pl present *currin*, which represents fourth conjugation morphology. It might be asked, however, whether the singular and 3pl present forms, and the full subjunctive, do not also belong to the fourth conjugation, meaning that there is no need to posit the N-pattern distribution suggested in Table 5.4 for *cuorrer* above.

With Maiden (2009: 84), I claim that this opposition within the present indicative does have morphological reality. One point to note is that verbs patterning with *cuorrer*, whether they are inherited or borrowed, systematically fail to show the fourth conjugation augment *-isch-*, which is otherwise productive in the fourth conjugation in Romansh. The augment is especially likely to be found in loanwords, and loanwords which do belong to the fourth conjugation across the paradigm (and thus take infinitive *-ir*) nearly all *do* display the augment; its absence from the relevant cells in the paradigm loanwords with an infinitive in *-er* is thus a reliable

sign that these cells are patterning with the third conjugation infinitive. What is more, on the very rare occasions that such forms do take an augment, it is not *-isch-* in Vallader, but *-esch-* (cf. *negliger* – 1sg present *neglig(esch)*, as listed in Tschärner 2012: 34), i.e. not the augment proper to the fourth conjugation in particular. We may conclude that the fourth conjugation affiliation proper to 1pl *currin* (etc.) is not shared by the N-pattern cells just mentioned.

It appears, then, that the pattern taken up by conjugational morphology across the paradigm of heteroclite verbs such as *cuorrer* in Vallader involves agreement in affiliation at least between the infinitive and the N-pattern cells. What is more, the diachronic generalization can be made that the infinitive is conservative in its behaviour: consider the infinitive of heteroclite *comparair*, which does not itself move over to the fourth conjugation, even though this shift would be predicted on the ground that the infinitive does not constitute one of the N-pattern cells in the second conjugation.

In diachronic terms, it is clear that the attested paradigm of inherited verbs such as *cuorrer* shows that they have undergone morphological developments analogous to those seen widely for Romanian in the previous chapter, a fact which in itself suggests that the paradigmatic split which Anderson treats as phonological is (at least *also*) morphological. However, there are also clear differences in the resulting paradigmatic distributions of conjugational morphology between the Romanian case and that seen here for Vallader. Importantly, it is notable that in the present case, there is no suggestion that what might be considered the PYTA part of the paradigm, consisting of the perfect and imperfect subjunctive (continuing the same formations as the Romanian preterite and pluperfect), remains aloof from the incursion of the innovative fourth conjugation morphology into the paradigm. This

disagreement with the behaviour seen in Romanian is natural, however, as we have already noted that in Romansh, unlike in Romanian, the existence of a distinctive stem form specific to forms descended from the Latin *perfectum* is extremely restricted lexically and plays little structural role in the paradigm. In the light of the Romanian facts, the behaviour of Romansh heteroclisis in this regard can be seen as negative evidence in favour of the same principle linking the diachronic spread of innovative morphology over the heteroclite paradigm to patterns of alternation which already characterize the inflectional system more generally.

Meanwhile, the past participle is not necessarily affected by fourth conjugation morphology, but can instead retain a distinctive stem, just as it does (unlike the perfect and imperfect subjunctive) in many second and third conjugation verbs in Romansh. However, the fact that the present participle is, like the past participle, recorded as retaining its third conjugation form alongside its innovative fourth conjugation counterpart perhaps suggests a role for the slight extraparadigmatic nature of both of these parts of the paradigm in conserving their existing morphology in spite of the influence of the incursive pattern; the reluctance of infinitive *comparair* is perhaps explained in this way too.

Particularly interesting is the behaviour of the future forms (1sg *currará* as opposed to **currirá*). Given the fact that the N-pattern distribution in Romansh does not include the future, we might expect the Vallader future tenses, in heteroclite verbs of the type seen here, to display innovative fourth conjugation affiliation. Their failure to show fourth conjugation forms even as variants, as was seen in the gerund and past participle, strongly suggests that the analogical influence of the infinitive has played a role, signalling that speakers recognize the close historical relationship between these two parts of the paradigm. At the same time, however, it should be

noted that in the verb system of Surmeiran, the analogical influence of the N-pattern does win out, giving e.g. 1sg *curriró* (rather than **curraró*) as the Surmeiran equivalent of Vallader *currará* (Signorell, Wüthrich-Grisch and Simeon 1987). The fact that these two different options exist, and were settled on by speakers of two closely related dialects in response to the same combination of analogical pressures, demonstrates something of the intricacy of the structural relationships underlying behind the verb paradigm and influencing its conjugational development.

5.3 From heteroclisis to new regularity? The case of Surselvan

As we have seen, a lexeme can be described as heteroclite in synchronic terms if the different cells in its paradigm bring together types of inflectional behaviour belonging to independently existing classes which, in other cases, must be taken as standing in opposition to one another. Thus the verb *cuorrer* in the Engadine varieties is synchronically heteroclite, and has been since our earliest records, because it inflects partly according to the third and partly according to the fourth conjugation, both inflectional classes which must be posited in any case in order to account for the behaviour of numerous other verbs. In the case of *cuorrer*, we know that this synchronic heteroclisis comes as the result of a process of diachronic incursion, in which the descendant of the third conjugation Latin verb *currere* at some point took on fourth conjugation forms in certain parts of the paradigm. On the other hand, this is not true of all such verbs in Engadine: it is very likely that at least some recent coinages with infinitives in *-er* never made use of the corresponding third conjugation finite forms in parts of the paradigm where fourth conjugation forms are now found.

Nonetheless, what makes it a relatively straightforward matter to identify synchronic heteroclisis in the case of Engadine *cuorrer*, and other verbs which behave similarly, is the continued robustness of the two conjugations whose behaviour it combines. According to all accounts of the Engadine verb system (but see caveats below), the third and fourth conjugations are well-populated in comparison with the number of *cuorrer*-like verbs that straddle them, which make up a rather small (though by no means negligible) group.

As discussed in the previous section, most presentations of the Rhenish varieties of Romansh likewise acknowledge the existence of equivalent groups of

verbs whose behaviour is midway between that of the fourth and third(/second) conjugations; they also grant them a similarly marginal role in the verb system. This goes for normative grammars and comprehensive descriptions of the standard varieties, whether aimed at native speakers or not (Thöni 1964, Signorell et al. 1987; Gregor 1982, Liver 1991, Haiman and Benincà 1992, Borodina 1969, and many more), but also dictionaries (Vieli and Decurtins 1962) historical presentations (Stimm and Linder 1989), dialectological research (Candrian 1900) and one thorough and theoretically elaborate description of the Surselvan verb system in its entirety (Tekavčić 1972/3 & 1974).

However, in the case of Surselvan, at least, this conception of the morphological behaviour in question here is somewhat idealized and out of date, because the innovative forms play a much more prominent role than has generally been recognized. This has important ramifications for the description of the conjugations in this branch of Romansh. In this section, I present a more representative picture of the Surselvan situation, following Winzap 1986, to my knowledge the first and only scholar to address this point, and associated literature. Furthermore, I show that the Surselvan facts illustrate one way in which diachronic heterocclisis can lead not to synchronic heterocclisis but to the existence of an entirely new pattern of regular inflectional behaviour, which in turn may affect the ‘hierarchy’ of inflectional classes as a whole. An intermediate stage in this process is marked by the widespread presence of a different non-canonical phenomenon: overabundance.

Winzap 1986 (the *Cudischet da verbs*), which provides the crucial data, is a manual of conjugation in standard Surselvan, including tables for around 1300 regular and irregular verbs, and as a result it is able to provide more substantial

information on the *lexical* distribution of fourth conjugation behaviour over verbs with third and second conjugation infinitives than is the case in less narrowly focused descriptions of the standard.¹³⁷ Like other scholars, Winzap recognizes the existence of a group of verbs for which this hybrid behaviour is found regularly and reliably: only fourth conjugation behaviour is available to these verbs in the relevant cells, and in his terms, ‘las fuormas primaras sin *-ein, -eis* fagessen mal las ureglias’ (p. 6).¹³⁸ The examples from this group given in his introduction, namely *assister*, *ceder* and *crescher*, are familiar from section 5.2 above, as they are all among those which are generally noted to display heteroclisis in descriptions from across Romansh. However, in contrast with the literature listed above for Surselvan and the other Rhenish dialects, the verb tables themselves suggest that in Surselvan this category (*verbs che vengan screts generalmein mo cun las finiziuns -in, -is*)¹³⁹ in fact contains no fewer than 80 verbs, including very many which show no sign of fourth conjugation morphology in any other Romansh variety (e.g. *batter*, *clauder*, *perschuader*, *manér*). What is more, fourth conjugation morphology is not limited to this already large group: for almost as many verbs again (around 70), both the ‘regular’ second/third conjugation and the fourth conjugation endings are listed as available, and Winzap comments, ‘Igl ei strusch pusseivel ed era buca necessari dad indicar quala fuorma che predomina’ (p. 6).¹⁴⁰ Examples include *beiber*, *vender* and *volver*.

¹³⁷ As concerns the *paradigmatic* distribution of fourth conjugation behaviour in these verbs, Winzap agrees with other authorities in noting that it affects not just 1pl and 2pl present indicatives but also the corresponding subjunctive and imperative forms. Unfortunately, as he is treating the written language, like most other scholars he devotes no attention to potential differences in pronunciation between open and close *e* (and thus to the possibility of identifying heteroclitite behaviour) in parts of the paradigm where the spelling of second/third and fourth conjugation endings is identical (e.g. imperfect <-*eva*>).

¹³⁸ ‘The original forms in *-ein -eis* [for 1pl, 2pl present indicative] would hurt the ears.’

¹³⁹ ‘Verbs which are generally written only with the endings *-in, -is*’

¹⁴⁰ ‘It is hardly possible and indeed it is not necessary to indicate which form predominates.’

Between them, these two groups account for the large majority of all the second and third conjugation verbs in Surselvan. In fact, among the *regular* verbs traditionally belonging to these conjugations, Winzap gives only two which are never found to take fourth conjugation endings (*moler*, *sedulér*), while Spescha 1989, whose presentation of Surselvan verb morphology simply fleshes out Winzap's own, adds two more (*setschecher*, *sepertschecher*). Meanwhile, two lexical families are listed by Winzap as preferring the original third conjugation endings in the written standard, namely *metter* and *rumper* and their prefixed counterparts (Spescha 1989 adds the family of *prender*); but he acknowledges that in everyday speech these can be heard with fourth conjugation endings too. With these few exceptions it is only highly irregular verbs, most of which are associated most closely with the second conjugation and thus could well have been implicated here too, that stand together in avoiding fourth conjugation morphology (*pudér* does not give e.g. **pudin* alongside, or in place of, genuine *pudein*).

The obvious upshot of these observations is that the pattern of behaviour apparently shown by relatively few verbs in Engadine, such as *cuorrer* and *insister*, has grown significantly in its distribution over the lexicon of Surselvan. In fact, there remains almost no territory left for it to expand into, and the acceptance of the innovative behaviour in the written standard suggests that wherever fourth conjugation morphology has become available to a given verb it is tending to overcome the previous norm. This is the interpretation provided by Winzap (1986: 5), who says of those verbs which have a choice between fourth and non-fourth conjugation endings: 'Il lungatg plidau dat la preferenza allas finiziuns dalla quarta conjugaziun. Il svilup dil lungatg va en quella direcziun.'¹⁴¹ Accordingly, written

¹⁴¹ 'The spoken language prefers the fourth conjugation endings. The development of the language is moving in that direction.'

style should privilege the non-fourth conjugation endings as far as is still possible, but even in writing this conservative principle should not be taken too far (*ibid.*).

This general observation is backed up by the more detailed findings provided in Winzap 1990, which reports on the research undertaken in the preparation of the *Cudischet da verbs*. The following table, taken from that report, gives an overview of the number of Surselvan speakers selecting each of the two possible 1pl present indicative endings corresponding to various infinitives, in response to a questionnaire given both to seminary students and to older linguistic informants in three regions of Surselva.

Table 5.5 Numbers of Surselvan speakers indicating a preference for each of the two available 1pl present indicative endings on selected verbs (Winzap 1990: 40, with modifications)

	informants (older)		students (younger)	
	<i>-ein</i>	<i>-in</i>	<i>-ein</i>	<i>-in</i>
<i>meder</i>	3	3	1	8
<i>pender</i>	2	4	2	8
<i>perver</i>	3	3	2	6
<i>porscher</i>	3	3	1	8
<i>preceder</i>	3	3	2	7
<i>prender</i>	2	4	2	7
<i>proceder</i>	3	3	1	8
<i>quescher</i>	3	3	2	6
<i>repeter</i>	3	3	0	8
<i>restrenscher</i>	3	3	1	8
<i>sclauder</i>	3	3	2	7
<i>spender</i>	2	4	0	8
<i>storscher</i>	3	3	2	7
<i>tender</i>	3	3	0	8
<i>tonscher</i>	3	3	0	8
total	42	48	18	112

It is unfortunate that Winzap does not provide a more detailed account of this survey and its methodology; he lists the findings of only a small portion of the complete questionnaire (though these are presumably representative of the whole), the numbers involved are relatively low, and little information is given about the

respondents. Nonetheless, although relatively informal in nature, as the only survey available which treats this issue it is valuable evidence for the modern state of affairs. If we can take these results as indicative of a linguistic change in progress¹⁴² as Winzap suggests, the trajectory involved is clear: all these verbs, without exception, are not only able to take on inflectional endings historically associated with the fourth conjugation, but are apparently abandoning those endings historically associated with the third and second conjugations in favour of the innovative forms.

The current state of affairs in Surselvan, then, appears to represent a substantial further development, still ongoing, which proceeds from the general heteroclite situation once common to all the Romansh dialects, as described in the previous section for Engadine and Old Surselvan and still presented as the Surselvan norm in e.g. Liver 1991. The essential point is that there is now no way to claim that fourth conjugation behaviour is available only to a limited number of the relevant verbs, most of which entered the language as learned borrowings. To my knowledge the same development has not taken place elsewhere in Romansh: recent manuals from other dialects (e.g. Tschärner 2012 on Vallader; Signorell et al. 1987 on Surmeiran) show no sign that fourth conjugation morphology is ever used by the general mass of inherited second and third conjugation verbs, such as reflexes of *manēre*, *bibere*, *volvere*.

This raises the question whether the Surselvan phenomenon came about as a result of particular dialect-specific pressures. However, I doubt that any such pressures are to be found. One possibility somewhat echoes the explanation proposed

¹⁴² Another conceivable explanation for the pattern seen here is age-grading: that is to say, rather than showing that the fourth conjugation endings are used more widely by more recent generations of speakers and thus that they are increasingly treated as the norm, these data could show a stable pattern whereby speakers in general use fourth conjugation endings on these verbs more frequently in their youth before gradually ‘growing out’ of the habit. However, this would not explain why the standard language now shows so many more verbs using the fourth conjugation endings than the Surselvan of previous generations.

for the heteroclite behaviour of Engadine *cuorrer* in particular in section 5.2 above. Arrhizotonic past participles from second and third conjugation verbs are marked with *-iu* in Surselvan (as in *tundiu* from *tunder*): this is likely to be a phonetic development from earlier *-ïiu* < *-ūtum* (cf. French *tondu*), but its identity with the inherited fourth conjugation *-iu* < *-ītum* could have been responsible for the spread of fourth conjugation morphology elsewhere in the paradigm of such verbs. However, the same innovative fourth conjugation morphology is also shown by verbs with rhizotonic participles (e.g. *clauder*, *claus*), and there is no evidence that verbs with participles in *-iu* have been affected more readily than others. Moreover, the same analogical change did not take place in Puter, even though the arrhizotonic ending *-i* in that dialect is already demonstrably borrowed from the fourth conjugation before our earliest records (as Puter, unlike Surselvan, does not show a sound change *ii* > *i*; cf. Eichenhofer 1999: 111). If anything, it may be worthwhile to suggest that Surselvan has apparently been more favourable than Engadine towards this particular pattern of heteroclisis for a long time: Mourin and de Poerck 1964 notes only one heteroclite verb family for OEng (that of *cuorrer*), but four for OSurs (*arver*, *scriver*, *viver*, *legier*). The Surselvan innovation of verbs of the type *cuvierer* with 1pl *cuvrin* (earlier Surselvan, and current Engadine, have *cuvrir*, *cuvrin*) also suggests that this heteroclite pattern enjoyed greater productivity long before the development discussed in the present section took hold. However, this is not intended as an explanation: it merely casts the same question further back into the past.

However this may be, what I wish to draw attention to in this section is not so much the spread of the innovative inflectional pattern over the lexicon as the shape of the overall morphological system which results, and the theoretical points which arise from this. What we see in Surselvan is the triumph of heteroclisis, in diachronic

terms. Practically no regular verb of the second or third conjugation ‘survives intact’ in the face of the specific process of incursion involving fourth conjugation forms, which is attested across Romansh. But in synchronic terms this ultimately has a paradoxical effect. What makes it useful to label Engadine *cuorrer* as heteroclite synchronically is the existence of two distinct and robust regular inflectional classes which it can be seen to straddle; it does not show any irregularity in its morphology except for the fact that it unexpectedly combines two *different models of regularity*. In Surselvan, by contrast, this is no longer the case. When Surselvan speakers produce 1pl *-in* on a second/third conjugation verb, they are no longer bringing together two distinct models of regularity at all: almost *all* verbs in those conjugations may inflect in this way, and most which cannot are already identifiably irregular in some other respect. Thus the current result of what might be called ‘diachronic heteroclisis’ in this case is not synchronic heteroclisis, as in Engadine: there is simply no group of verbs left in Surselvan to constitute the ‘regular’ type which in Engadine is still embodied by the more conservative descendants of the Latin second and third conjugation verbs. The equivalents of the Engadine heteroclites are now themselves regular verbs in Surselvan.

At present it would still be premature to say that the endings which historically belong to the fourth conjugation (1pl *-in* etc.) have altogether lost this association. This is because verbs in the fourth conjugation can only take those endings, and there is no reason to think that speakers vacillate at all in the way they inflect those verbs in the relevant cells. Meanwhile, as shown impressionistically at least by the table in Winzap 1990, verbs which historically belong to the second and third conjugations do still retain the possibility of using the more original morphology (*-ein* etc.) among at least some speakers. In fact, it is notable that the

figures provided by Winzap do not bear out the distinction he makes between verbs accepting fourth conjugation morphology to different extents: for example, according to his data, *proceder* and *preceder* behave similarly to the other verbs in question.

It seems, in fact, that the most appropriate way to discuss the behaviour of the verbs which have undergone this morphological shift is to say that they show systematic *overabundance* in the sense of Thornton 2011, offering both conservative second/third conjugation and innovative fourth conjugation forms: both *-in* and *-ein* (etc.) are available to speakers as regular forms for practically all the verbs of Surselvan belonging historically to the second and third conjugations, and accordingly they differ both from fourth conjugation verbs, which may only take *-in*, and first conjugation and irregular second/third conjugation verbs, which may only take *-ein*. Interestingly, this overabundance is not merely present at the level of the variety as a whole but can also be detected at the level of individual speakers, as Winzap records that his informants sometimes admitted to uncertainty about which form they themselves used. Winzap's study thus offers us a valuable snapshot of conjugational incursion in progress.

5.4 The infinitive in Val Müstair

The final development to be discussed in this chapter is restricted to a very small portion of the Romansh-speaking domain. Val Müstair (VMü), a valley lying at the eastern extreme of the canton of Graubünden and remote from the rest of the Romansh zone, is home to just over 1000 of Switzerland's 40,000-60,000 Romansh speakers, and as a dialect area comprises only six of the 148 settlements which serve as points of enquiry for the dialectological survey of the language (listed as E30-E35 in the coding used in the *Dicziunari Rumantsch Grischun* and elsewhere). The most populous of these settlements, with around 600 inhabitants, and the source of most of the data used in this section, is the village of Müstair itself (E35) on the Italian border; its local variety is treated by Schorta 1938, a monograph concentrating largely on questions of historical phonology. Though the issues treated in the present section are also touched on in Eichenhofer 1999, Maiden 2011b, and Anderson 2013, that monograph is still the only work devoted to the language of Val Müstair in particular (and provides the linguistic material relayed here unless otherwise stated), and it draws attention to a grammatical phenomenon which, I will argue, deserves to be viewed diachronically as an instance of the partial conjugational change which has been seen to give rise to heteroclisis in some of the cases treated in this thesis. On the other hand, while the precise conditions on the lexical distribution of this change are far from clear, it is harder to see its results as constituting heteroclisis in the *synchronic* inflectional system of Val Müstair. To this extent, there is an obvious parallel between the development investigated here and that seen for Surselvan in the previous section, where conjugational incursion has likewise resulted not in heteroclisis but in an innovative type of regular behaviour. However, the specifics of the case are very different; what is more, alongside its own inherent interest, the

investigation of this development touches on a wider historical question concerning the nature of the first conjugation augment as it appears in Romansh.

Unfortunately, Schorta goes into little detail on the phenomenon in question here, and the dearth of information available on this and the morphology of VMü in general means that my analysis here is inevitably provisional. Nonetheless, it does point to a morphological development which is without parallel elsewhere in Romansh, and which seems to show an origin for conjugational change otherwise unattested in the history of the languages treated in this thesis. Here an inflectional development affecting the distribution of conjugational morphology over the lexicon has apparently arisen as a result of stress retraction, originally a matter of post-lexical phonology, which came to be morphologized as a feature of the infinitive in certain verbs.

As indicated by the coding assigned to them by *DRG*, the varieties spoken in Val Müstair belong to the Engadine dialect group, and more specifically to Vallader. However, within Vallader the VMü varieties form a divergent cluster of their own, to the extent that some prefer to treat so-called *Jauer* (from VMü *jau* ‘I’, cf. standard Vallader *eu*) as a third Engadine dialect distinct from both Vallader and Puter (Liver 1999: 42f.). This situation makes sense geographically, as Val Müstair is not in fact part of the Engadine valley proper but is separated from it by some miles of uninhabited mountainous terrain; access from the Engadine is only possible by way of the Fuorn Pass, one of the highest passes in the Swiss Alps. The distinctive characteristics of VMü phonological development which this isolation has engendered are documented by Schorta 1938 (see especially pp. 133-35 for an overview): these include both idiosyncratic innovations (e.g. *nct* > *ñć*, as opposed to *nct* > *t* as seen elsewhere in Vallader) and conservative features found elsewhere in

Romansh but not retained by the rest of Vallader (e.g. *j* remains after labial consonants, as in Surselvan at the other end of the Romansh dialect continuum, rather than undergoing fortition to *ǵ*).

However, as Schorta recognizes (p. 132), by far the most striking and unusual feature of the Romansh of Val Müstair, found nowhere else in Romansh¹⁴³ or indeed the whole Romance world, is not any such outcome of ‘blind’ historical phonology; rather, it is a phenomenon limited to the inflectional behaviour of the verb. The nature of the phenomenon is illustrated in the following table, in which the Val Müstair forms are given in a phonemic version for expository purposes:

Table 5.6 Selected verb forms in Vallader and VMü

Vallader infinitive	Val Müstair infinitive	Val Müstair 3sg pres. indic.	
<i>arár</i>	'arər	'arə	plough
<i>ladár</i>	'lajdər	'lajdə	fertilize with manure
<i>filár</i>	'filər	'filə	spin
<i>telefonár</i>	'tele'fonər	'tele'fonə	telephone
<i>fundár</i>	'fondər	'fondə	melt
<i>buglír</i>	'bwɔ́lər	'bwɔ́lə	boil
<i>dormír</i>	'dɔ́rmər	'dɔ́rmə	sleep
<i>sesáir</i>	'setsər	'setsə	sit

(Schorta 1938: 132; Eichenhofer 1999: 41f.; *HWR passim*)

The remarkable feature in question is captured in the difference between the forms given in the first two columns, which contrast a range of VMü infinitives with their counterparts elsewhere in Vallader (the relevance of the third column will be made clear later). For all the verbs presented in this table, the infinitive found in the other Vallader varieties, exemplified here by the form used in standard written Vallader, is affiliated distinctively with the first, second or fourth conjugation, as marked by the characteristic stressed infinitive endings (*-ár*, *-áir*, *-ír* respectively). In the

¹⁴³ ‘Die auffallendste Eigentümlichkeit der Mundart von VMüst. ist zweifellos die Zurückziehung des Akzentes beim Infinitiv der Verba... Diese Eigentümlichkeit trennt die Mundart von VMüst. scharf von allen anderen rätoroman. Mundarten.’ (Schorta 1938: 132).

corresponding VMü forms, however, this is not the case: in place of the various stressed endings which would be expected on etymological grounds, these all display the unstressed ending *-ər*, which is identical with the inherited marker found in VMü, as throughout Romansh, on regular infinitives of the *third* conjugation (e.g. *'mɛt-ər* < *mett-ere*). It seems likely that the change which gave rise to these forms is a fairly recent one: Schorta 1938 reports the existence of the newer infinitive *'jajntər* ‘eat lunch’ alongside a more conservative, and distinctively first conjugation, *jən'tar* < *ientāre*.

However, the situation in VMü is not one of total neutralization of inflectional class distinctions in the infinitive. On the contrary, all three stressed infinitive endings seen elsewhere in Vallader do survive in VMü, e.g. in first conjugation *bə'tjar* ‘baptize’, second conjugation *vø'ɫajr* ‘want’, and fourth conjugation *fi'nir* ‘finish’, none of which show signs of being replaced by forms in *-ər*. But the apparent use of third conjugation morphology in the infinitive of verbs to which it does not historically belong is extremely widespread, and by no means specific to the verbs given as examples in the table above, which merely happen to be among those for which the relevant forms are explicitly provided in the literature. In fact, use of the *-ər* ending can now be taken as the norm, while those verbs which do retain the stressed endings in the infinitive are the exceptions requiring a special account, as discussed below: this has important ramifications for the synchronic analysis of the conjugational system as a whole. It is notable, though, that we are not dealing with the spread of third conjugation membership as such across the lexicon; a given lexeme’s adoption of the infinitive in *-ər* does not signal the wholesale transfer of that lexeme into the third conjugation, complete with the adoption of the appropriate morphology throughout the paradigm. On the face of it we are thus

dealing with conjugational incursion, whereby third conjugation morphology has infiltrated the paradigm of verbs originally belonging to other conjugations but has got only so far.

The introductory presentation given in the previous paragraphs takes for granted that the aberrant behaviour of the VMü infinitive is a morphological issue, involving the replacement of one conjugation's morphology with another's. But this deserves further comment. As can be seen in the examples given above, the obvious corollary of the adoption of an unstressed ending (*-ər*) in place of a stressed ending (such as *-'ar*) is that word stress in the innovative forms is retracted by one syllable, on to the final syllable of the stem; and each of these developments is only possible in the presence of the other. Accordingly, if it makes any sense to distinguish between the two in synchronic terms, it should not be taken for granted that the alteration undergone by the infinitive ending is responsible for the prosodic shift rather than vice versa. This raises the possibility that it is wrong to see any interesting morphological, as opposed to phonological, development here at all. That is to say, the existence of an infinitive form *'arər* (for standard Vallader *arár*) might not demonstrate that the morphological behaviour of this verb in that particular cell should be identified with the behaviour characterizing the third conjugation; it might instead result from a relatively superficial phonological phenomenon of stress retraction acting, for whatever reason, on an underlying infinitive *ər'ar*, identical to that seen in standard Vallader and belonging to the first conjugation as expected. Indeed, in the relevant sections of Schorta 1938 and Eichenhofer 1999, both manuals on historical sound change, it is the stress retraction that is taken as primary, while the change in the infinitive ending is treated simply as a side effect: neither scholar

even points out explicitly that the *-ər* found in these verbs can be likened to the inherited third conjugation infinitive ending.¹⁴⁴

As far as the historical direction of causation is concerned, at least, this is in fact likely to be the right way round. Accounting for the unusual behaviour of VMü infinitives, Eichenhofer points out (pp. 41f.) that verb forms in Romansh are often found in combination with a following adverb: as examples involving the infinitive, he provides reconstructed VMü syntagms such as **yz'ar 'jo* ‘wear out’ (literally ‘wear down’, corresponding to standard Vallader *usar jo*).¹⁴⁵ In cases where the verb involved belongs to the first, second or fourth conjugation, the lexical stress which is expected to be present on the infinitive ending thus clashes prosodically with the stronger sentence stress assigned to the adverb, and as a result undergoes retraction to the preceding syllable (cf. the well-known ‘rhythm rule’ seen in English examples such as *Heath'row* vs. *'Heathrow 'Airport*, *fif'teen* vs. *'fifteen 'minutes*: Liberman and Prince 1977: 309f.). He claims that the original infinitive endings would automatically be reduced to *-ər* in this position; the generalization of this originally context-dependent stress pattern to all syntactic environments would thus give rise to the phenomenon seen in VMü.

Although Eichenhofer provides no indication that independent features of VMü can be identified¹⁴⁶ which left that particular variety especially liable to make

¹⁴⁴ However, Schorta (1938: 32) does recognize that the morphology is implicated in this development, presumably on the basis of the considerations I point to in the following paragraphs.

¹⁴⁵ Another of his examples, **sə'tsajr 'jo* ‘sit down’, intended as the precursor of current *'setsər 'jo*, may be irrelevant, as this verb has moved to the third conjugation (i.e. has taken on a rhizotonic infinitive, in contrast with Latin *sedēre*) not just in VMü, but widely across Romansh. However, as Eichenhofer notes, this could be precisely because in this case the prosodic effect proposed here became morphologized outside VMü as well.

¹⁴⁶ It is hard to make suggestions in this area, given the scarcity of available information on any distinctive prosodic or syntactic characteristics which VMü may possess. One relevant factor can be pointed to, however. If it is right to claim, as I do, that the generalization was made possible because speakers identify the de-stressed *-ər* found in these innovative infinitive forms with the inherited third conjugation marker, it is at least no surprise to find that this has taken place in an eastern dialect of Romansh (= Engadine or Surmeiran) rather than a western dialect (= Surselvan or Sutselvan), as in the

this generalization – on the basis of a prosodic phenomenon which presumably applies across Romansh – this is nonetheless a plausible account of the origins of the idiosyncratic development seen in VMü, and one which I tentatively support. Nonetheless, the likelihood that the initial trigger for this behaviour was indeed an automatic prosodic adjustment of this kind does not mean that the current state of affairs in VMü can be explained in purely phonetic or phonological terms, or that the aberrant VMü infinitives should synchronically be thought of merely as surface variants of forms which somehow underlyingly show the same conjugational behaviour as is seen elsewhere in Vallader.

Instead, it is clear that we are dealing with a phenomenon which, whatever its origins, has become articulated by the conjugational system in various ways. For one thing, the morphological nature of the development is implicit in the statement that the use of *-ər* in the infinitive of the relevant verbs is no longer limited, if it ever was, to any particular prosodic context, while the same thing is not reported to have taken place in any other verb form (or indeed anywhere else in the language at all). Even to the extent that it makes sense for speakers to continue characterizing the appearance of *-ər* in etymologically non-third conjugation verbs as the result of superficial stress retraction – which arguably remains one potential interpretation for as long as rival old and new forms such as *jən'tar* and *'jajntər* coexist(ed) in the language – the fact that it now applies even in the absence of a following sentence stress means they must have reinterpreted this retraction as conditioned by the specific inflectional identity of the form as a verb in the infinitive.

Moreover, if the diachronic explanation adopted here holds water, the very fact that the use of *-ər* spread to contexts where it was not prosodically motivated

latter the stressed infinitive endings have lost their final *-r* and therefore in a prosodically weakened form would never be identified with the *-ər* of the third conjugation.

calls out for an account which recognizes the contribution made by speakers' morphological knowledge. In ordinary circumstances it would be hard to explain why speakers might have been tempted to generalize to all contexts a variant unstressed ending which started off as the entirely predictable result of post-lexical stress clash, and only appeared in the presence of an unambiguous conditioning environment – especially since instances of the basic, stressed ending must have been so commonplace for all but those verbs which were never found without a following adverb (if any even existed). The prior existence of a class of verbs which already marked the infinitive with an invariable unstressed *-ər*, namely the third conjugation, is likely to have been crucial in making this otherwise unexpected shift possible. At the very least, the success of the generalization of 'secondary' *-ər* to all contexts must have been favoured by the example of the third conjugation verbs, some of them among the most frequently used in the language (e.g. *'metər*, *'prendər* and their compounds), for which *-ər* was the ending found in *all* prosodic contexts.¹⁴⁷

In any case, treating the development seen here as a question of the lexical generalization of *-ər* is no more than a convenient shorthand, as the behaviour of the infinitive marker itself in these forms is only half the story. As mentioned above, the unstressed ending *-ər* is not the only distinctive characteristic of the VMü infinitives treated here: logically, the adoption of this ending goes hand in hand with the

¹⁴⁷ Note that in order to propose, as I do here, that the *-ər* seen in VMü infinitives of verbs which do not etymologically belong to the third conjugation arose as a de-stressed variant form, before later becoming basic when speakers identified it with the existing third conjugation infinitive ending, we do not need to be certain that *-ər* would be the expected automatic reflex of all three stressed infinitive endings (*-'ar*, *-'ajr*, *-'ir*) under effacement of the stress. This is apparently assumed to be the case by Eichenhofer; but while *-ər* is phonetically plausible as a weakened version of *-'ar*, it is not obvious that, say, de-stressed *-'ir* 'should' have given *-ər* directly rather than an ending containing a front vowel. It could well be that *-ər* as a prosodic variant was originally specific to verbs in *-'ar* alone; in that case, it must have been reanalysed initially as a morphologically novel infinitive ending available to (certain) first conjugation verbs, and only later generalized to the second and fourth conjugations. To weigh up the relative likelihood of these developments would be difficult given the state of our data on VMü, and would take us too far afield here.

innovative placement of stress on the stem. And in fact the shape that the stem takes on as a consequence provides a further indication that, whatever its origins, the behaviour of these innovative infinitives cannot be treated synchronically as the result of a ‘blind’ process of stress retraction without reference to the morphological system.

This is because of the way in which the formation of these infinitives interacts with a central feature of the Romansh verb system discussed in the previous section. In VMü, as in other Romansh varieties, the great majority of verbs build their inflected forms on the basis of two phonetically distinct stems, whose distribution over the paradigm is correlated with the distribution of stress; this is the Romansh instantiation of the morphomic N-pattern found widely across Romance. Forms of the verb ‘to sing’ provided in Schorta 1938 (pp. 80 and 126 respectively) exemplify this alternation within the paradigm: 2sg present indicative *'cawnt-əš* (< *cánt-ās*), showing the stem found when stress is located on the stem itself, as opposed to m. sg. past participle *cənt-'a* (< *cant-átum*) in which stress is located elsewhere. However, as mentioned above (with the works cited there) with reference to other Romansh varieties, although the two stems of any given verb are predictably associated with opposing stress conditions, they do not stand in a similarly predictable phonological relationship with each other. Both are usually reflexes of a single Latin stem form, but the workings of differential sound change from this common starting point have given rise to a large number of alternation patterns over the lexicon as a whole; these cannot be accounted for in terms of synchronic phonology, and must instead be learnt as idiosyncratic properties of individual lexemes.

It is true that, to a greater extent than in the Rhenish varieties of Romansh, which preserve scores of these alternations as they were produced by regular sound change (cf. Anderson 2011: 19-22), the Engadine varieties have seen processes of analogical levelling reduce the number of alternations available and restrict the disparity between paired stems. For example, unlike Rhenish, Engadine does not possess pairs in which the alternation affects more than one vowel of the stem: compare Surselvan *dumandár* ‘to ask’ – *damónd* ‘I ask’ and the slightly remodelled Vallader *dumandár* – *dumónd*. However, even in Engadine, and in VMü specifically, the phonological unpredictability of the relationship remains. Crucially for the present topic, knowledge of a verb’s unstressed or non-N-pattern stem, as is found in conservative non-third conjugation infinitives, is not always sufficient to determine its stressed or N-pattern stem: for example, the alternation of (unstressed) *ə* with (stressed) *aw*, seen above in *cənt-* vs. *cawnt-* ‘sing’, is not matched by other lexemes such as *sər-* vs. *sər-* ‘shut’ or *cən-* vs. *cajn-* ‘dine’.

What this means is that the retraction of stress alone cannot be directly responsible for the stem form found in the innovative infinitives of VMü. As far as the ending *-ər* seen on these infinitives is concerned, it is at least conceivable (although unlikely for the reasons given above) that it arose purely as a phonologized version of the weakened vowel which resulted from basic *-ár/ájr/ír* with the removal of the lexical stress, and that speakers have never identified it with the inherited third conjugation marker; that is to say, it might have developed without undergoing any influence from the rest of the verb system as this was already set up. But something besides the retraction of stress is required to explain the vocalism of the newly stressed *stem* of these infinitives, which is not predictable from the vocalism of the unstressed stem which they previously displayed (compare the infinitives

corresponding to *cən-*, *sər-* and *cənt-*). In fact, as Table 5.6 clearly shows, the new VMü infinitives systematically use the same stem as is already found in the N-pattern section of the paradigm (illustrated here by the 3sg present indicative form): this is why inherited *ərar*, *lədar* give way to *arər*, *lajdər*, and not (for example) **əjrər*, **lədər*, respectively. Thus even on the assumption that the original impetus for the development of these distinctive VMü infinitives belonged to the prosody, the stem shape they now show did not come about in isolation, but must be the result of morphological analogy with the existing N-pattern forms of each given verb.¹⁴⁸

Another way to put this is that speakers have incorporated these infinitives into the N-pattern – thereby giving them a distribution of stems over the paradigm which already featured in the verb system, but was previously a distinctive marker of third conjugation verbs. The behaviour of the stem thus confirms the interpretation proposed above of *-ər* as a *morphological* interloper; the innovative VMü infinitives are not just coincidentally like third conjugation infinitives, but their morphological composition *is* that of third conjugation infinitives, in the sense that they have taken on morphological behaviour that previously characterized the third conjugation exclusively.

In terms of the history of the verb system, then, we can be confident that in this characteristic phenomenon of the Val Müstair variety we are dealing with what can be called the diachronic counterpart to heteroclysis, namely the incursion of inflectional morphology from one class into the paradigm of lexemes outside that class. And while not all of the developments treated in the course of this thesis have

¹⁴⁸ The analogical pressure can also apply in the other direction, from the new infinitive to the other N-pattern cells. For example, the sequence *-ld-* seen in the VMü verb for ‘hear’, infinitive *'daldər* (3sg *'daldə*) from **ad-audire*, would be the expected outcome only in pre-tonic position. Its presence in the 3sg and other N-pattern cells is likely to be due to the fact that it was retained in the infinitive when this took on its new stress pattern. In any case, this would still demonstrate that the infinitive had joined the N-pattern. (On the other hand, it cannot be ruled out that in this case the analogy had already taken place before the infinitive changed its stress pattern.)

necessarily involved the innovative adoption of extraneous inflectional morphology in this way – recall from section 3.2 above that the proximate cause of the Latin *capiō* type’s behaviour seems to have been neutralizing sound change, in combination with related sound changes affecting what would become the third and fourth conjugations, with no further morphological ‘interference’ required – such incursion is a major source of synchronic heteroclisis, and we have been able to detect the lexical spread of heteroclisis arising on this basis in all the languages treated here. Within Romansh specifically, the type of heteroclisis treated in the first section of this chapter, involving the encroachment of fourth conjugation morphology into the second and third conjugations, was treated as a characteristic of particular lexemes: in Engadine and at least in early Surselvan, particular verbs can be identified which have taken on unexpected fourth conjugation forms, and they represent lexical exceptions to normal conjugational behaviour. And although I have argued that, in Surselvan, what was previously the exception has now become the norm, the spread of this new norm has been (and apparently remains) lexically gradual: we do not find that some inherited second/third conjugation verbs take on fourth conjugation morphology in certain parts of their paradigm, and others hold out against it, in accordance with any formal or semantic characteristic that can be seen to license the change.

Like fourth conjugation morphology in Surselvan, the VMü infinitives in *-ər* are found widely beyond their point of origin, to the extent that they can apparently now be taken as the norm from which certain exceptional verbs diverge. Unfortunately, as VMü is fundamentally an unwritten variety, we know much less than in the former case about how the innovative forms came to acquire this predominance over the verb lexicon. But one synchronic difference between the two

cases is clear: it is generally agreed, albeit by the relatively few scholars that have addressed this question, that a given VMü verb's use of or failure to use *-ər* is better taken not as an inherently unpredictable fact about that verb, but as governed by certain principles. On the other hand, exactly how those principles should be captured is still an open question, as I will show. Their investigation also brings up a historical question about Romansh verb morphology which is of interest in its own right.

The point which proves central to the discussion is brought up for the first time, to my knowledge, by Eichenhofer (1999: 41). Although he initially describes the unexpected appearance of *-ər* in the infinitive in VMü as primarily a characteristic of *first* conjugation verbs (as does Schorta 1938: 132), he does acknowledge that the phenomenon affects some verbs belonging to the second and fourth conjugations as well. But, as he points out, fourth conjugation verbs are not affected if their present tense paradigm employs the productive 'augment' descended from *-isc/esc-*, whose use characterizes the majority of fourth conjugation verbs in Romansh: thus *finir* 'end' (3sg *finišə*, DRG 6.346) and *pətir* 'suffer' (3sg *pətəšə*, HWR 2 s.v. *pitir*)¹⁴⁹ have evidently resisted the change that has taken place in verbs such as *'dormər*, *'bwəlar* which do not feature the augment.¹⁵⁰

One potential interpretation of this finding is that the behaviour of the VMü infinitive can be constrained by the morphological structure found in other forms in the paradigm. This would point to an account which assumes restrictions on the paradigmatic distribution of the augment. Given the fact that, where innovative third conjugation infinitives are found in VMü, they reliably pattern with N-pattern forms

¹⁴⁹ It is unclear why the augment is recorded as taking the form *-iš-* in one case and *-éš-* in the other. I refer to it in general as *-iš-* in the subsequent discussion.

¹⁵⁰ This explains why the development is associated more closely with the first conjugation than the fourth, as most fourth conjugation verbs do use the augment and are accordingly barred from showing this innovation in the infinitive.

such as the 3sg present (as seen in Table 5.6 above), one might expect that if – contrary to fact – the form *finišə* were to be accompanied by an infinitive in *-ər*, this would take the form **finišər*. The fact that speakers do not make use of this possibility may then suggest that they treat the augment as simply morphologically incompatible with the infinitive ending *-ər*, potentially because the distribution of the inflectional augment over the paradigm in the inherited Romansh system never does include the infinitive cell, or indeed any non-finite cell: this may just be a contingent fact, or it may be treated as the result of some stipulation (Anderson 2013: 12, 20f.), but either way, there is no existing Romansh model promoting the creation of such infinitive forms in VMü.

This is the approach taken by Maiden (2011b: 45f.), in addressing the failure of some verbs belonging to the *first* conjugation to display the new infinitive in *-ər*. Maiden claims that the first conjugation in VMü contains verbs making use of a *distinct* inflectional augment, with the shape *-aj-* (for example 3sg *guj-áj-ə* ‘dares’, *bəťj-áj-ə* ‘baptizes’), and these are systematic in retaining the inherited infinitive form (*guj-ár*, *bəťj-ár*) because this augment is restricted morphologically from appearing in the infinitive and forms such as **gujájər* thus could not exist.

Although Maiden does not mention Eichenhofer’s observation, the parallel with the behaviour seen in the fourth conjugation is striking. What is more, Maiden’s claim, in using the term ‘augment’ with reference to the stem alternation seen here in *bəťjájə/bəťjár*, is precisely that the *-aj-* element seen in these verbs is a distinct first conjugation counterpart to the augment seen in the fourth conjugation.¹⁵¹ This is clear from the context of the discussion, where the behaviour of the fourth

¹⁵¹ Thus, Maiden’s observation (2004: 24) that ‘[w]hereas in most Rhaeto-Romance varieties the first conjugation augment *-edj-* is maintained, in Engadine and Surselvan Romansch dialects it has been systematically substituted by the fourth conjugation augment *-ef-*’ must be taken to count the VMü varieties as distinct from Engadine (while a part of Rhaeto-Romance more widely).

conjugation augment *-esch-* is otherwise at issue, and presupposes (Maiden, p.c.) that Val Müstair, uniquely for contemporary Romansh, preserves an earlier pan-Romansh state of affairs in which the fourth and first conjugations made use of productive augments from different sources – the fourth conjugation *-iš-* reflecting the suffix *-ēsc/īsc-* as discussed above (section 5.1), and the first conjugation *-aj-* continuing a suffix which also survives as an inflectional augment in several other Romance varieties, namely the Greek verbalizer *-iz-* borrowed into early Romance in a shape generally reconstructed as **-idj-*.

This suffix **-idj-* is in fact the source of the productive Romanian first conjugation augment *-ez-* encountered in the previous chapter, which acts entirely in parallel with the productive fourth conjugation augment *-esc-* (Zafiu 2013: 25f.). Somewhat closer to home for present purposes is Dolomitic Ladin, where the reflex of **-idj-* behaves in just the same way, appearing in the N-pattern cells but not otherwise: for example (in the Ladin of Fodom, Pellegrini 1974: 43) it is absent from the infinitive *brontol-é* ‘grumble’ < *-āre*, but present (in the form *-ei-*) in 3sg present *brontol-ei-a*, as if < **-idj-at*. What is more, it is straightforward to recognize the presence of this same suffix in *bəťjájə* itself (< **baptidjat*) on etymological grounds. There is also no doubt that the alternation displayed by *bəťjár/bəťjájə* is, or has been, productive in the formation of new vocabulary, as shown e.g. by *məňjár – məňjájə* ‘think (i.e. have the opinion)’, with a stem borrowed from German *meinen* ‘think, mean’ (*HWR* 1 s.v. *maniar*).

Given that verbs such as *bəťj-/bəťjáj-* seem to contain a productive, isolable N-pattern stem marker (like *fin-/finiš-*), and moreover that elsewhere in the Romance world (including elsewhere in ‘Rhaeto-Romance’) cognates of the suffix involved are themselves employed as first conjugation augments sharing their paradigmatic

distribution with those of the fourth conjugation, it is not hard to see why it might be attractive to identify the *-aj-* of *bəťjájə* as historically representing an augment itself. However, it should be noted that Maiden is apparently practically alone in making this claim, which is not shared by Schorta, Eichenhofer, or indeed Meul (2013: 193), who in her dedicated survey of Romance inflectional augments states outright for Romansh that ‘no trace of *-idj-* is to be found in these varieties’. Maiden’s position is otherwise held, in fact, only by Anderson (2013: 20f.), who is responding directly to Maiden 2011b and presumably following the latter’s lead on this point, which is somewhat tangential to his own concerns.

And in fact, despite the synchronic similarities, we can certainly say that the formal relationship between *bəťjár* and *bəťjájə* in this specific verb did not arise historically in the same way as that between Ladin *brontolé* and *brontoleia*, or indeed between VMü *finír* and *finišə*. As mentioned above (section 4.1), the inflectional pattern illustrated by the latter comes as the upshot of an early Romance phenomenon which saw the falling together of paired lexemes which differed only with regard to the presence of the originally derivational element *-ēsc/īsc-*; this element was reinterpreted as an empty piece of morphological form required in certain parts of the paradigm of a large and productive class of verbs. The first conjugation augment, where it is found, results from the early reinterpretation of **-idj-* along the same lines (Meul 2013: 74f.).

By contrast, evidence from the other varieties of Romansh and from VMü itself shows that *bəťjár – bəťjájə* does not reflect the collapse of two lexemes into one in this way. Within VMü, alongside what is ostensibly *bəťj-áj-ə*, we also find attested the form *bəťáj-ə* (DRG 2.259) which does not lend itself to further segmentation as it does not contain the sequence *bəťj-* found in the infinitive. This latter form can be

taken as the more conservative, given that it is also the one found in standard Vallader (written *battája* alongside infinitive *battiár*); further afield in Surselvan, the corresponding forms are *battégia*, *battegiár* (for all these see *HWR* 1 s.v. *battiar*). Crucially, however, all the attested dialectal counterparts of VMü *bəťjár* agree with each other in reflecting regular developments from a verb **baptidjäre – *baptidjat* in which the element **-idj-* taken by Maiden as an inflectional augment is in fact present throughout, and simply represents part of the lexical material of the verb. There is no reason to mark out the VMü situation as historically separate from the rest: rather, as in standard Vallader,¹⁵² this lexical **-idj-* is reflected in the yod of the VMü infinitive *bəťjár*. N-pattern forms of the type *bəťjájə*, as opposed to conservative *bəťájə*, simply reflect the analogical insertion of this yod on the basis of forms such as *bəťjár* at a recent stage.

In diachronic terms, then, this verb belongs not with the augmented verbs of VMü but with the ‘alternating verbs’ (cf. Table 5.1), like its counterparts in all other varieties of Romansh. To take a broader Romansh perspective, it does not historically share a formation with Ladin *brontol-é brontol-éi-a* or Romanian *lucr-á lucr-eáz-ă*; instead, the appropriate comparison is with situations in other languages where **-idj-* is treated as a derivational marker (French *-oyer*, Italian *-eggiare/-izzare* etc.: Maiden 1995: 189) or a fossilized part of the lexical stem. Most strikingly, its historical formation corresponds precisely to that of its cognate in Romanian, *botez-á* (3sg *boteáz-ă*) ‘baptize’, one of the few Romanian verbs in which *-ez-* < **-idj-* is not an inflectional augment but part of the lexical material.

I see no reason to doubt that all attested instances (and in fact these are only few altogether) of an apparent inflectional augment *-aj-* in VMü are in fact to be

¹⁵² A similar phonetic outcome is seen in the Surselvan dialect area of Cadì, which has infinitive *bəťiá* (Eichenhofer 1999: 390).

explained in the same way. The inheritance of **-idj-* as a derivational element is sufficient to account for the existence of productively derived verbs such as *mənjár – mənjájə* ‘think’ as observed above, and there is apparently no evidence that inherited first conjugation verbs are liable to acquire an productive *-aj-* element as an inflectional characteristic in the way that *-esch-* can be taken on by first conjugation verbs in other Romansh varieties (cf. Vallader *speresch* ‘I hope’ reflecting Latin *sperō -āre*, as noted in section 5.2). Meanwhile, Vallader (along with the other varieties) contains numerous verbs whose inflectional alternation across the paradigm is clearly akin to that of VMü *bəjtár*, without thereby being taken to continue the inflectional augment (e.g. *nettiar – nettaja* ‘clean’ and indeed *maniar – manaja* ‘think’, cf. *HWR* for both). I claim, then, that Maiden is wrong to see the inflectional behaviour of verbs such as *bəjtár* as evidence for the preservation of an earlier inflectional augment derived from **-idj-* < Greek *-iz-*, which was originally characteristic of productive first conjugation verbs throughout Romansh and has since been replaced in other dialects by *-esch-* from the fourth conjugation within the course of the language’s history as a distinct entity.¹⁵³

To return to the specific question of the lexical distribution of the infinitive in *-ər*, however, it must be said that the differences in the *historical* reasons for the paradigmatic alternations seen in *finir – finišə* and *bəjtár – bət(j)ájə* are not necessarily germane to a synchronic morphological account. It could be that speakers synchronically treat the two as having the same morphological status, on the grounds that they both show a difference in syllable count: the very fact that *j* has come to be

¹⁵³ This may nonetheless have been the case, given the existence of a genuine inflectional augment < **-idj-* in Ladin: I observe simply that *bəjtár* and the like provide no good evidence for this possibility. It must also be noted that Schorta 1938 contains no attestations of a first conjugation augment *-esch-* in VMü; however, this is not necessarily unexpected, as it is not primarily concerned with morphology, and the augment is less commonly encountered in the first conjugation than in the fourth throughout Romansh.

inserted secondarily to give forms such as *bətjájə* is perhaps an indication that *-aj-* has come to be viewed as parallel to *-iš-* in this way. Maiden's account of this particular phenomenon is therefore not to be ruled out.

However, it might be equally appropriate to treat the fact of this difference in syllable count as being *directly* in question here, rather than because of what it implies to speakers about the morphological nature of the 'extra' material. The reason why (originally automatic) stress retraction on an inherited infinitive such as *arár* 'plough' or even *ladár* 'fertilize' could be reinterpreted as morphologically relevant is evidently related to the existence of a certain level of resemblance in shape between the stress-retracted versions of these infinitives and the existing N-pattern forms such as 3sg *árə*, *ládə*. In the case of *finír* alongside 3sg *finišə*, however, the similarity was much less; the same applies in the case of *bəjár* alongside 3sg *bət(j)ájə*, regardless of the morphological underpinnings of these forms. This is to say that it could be the level of *phonological* similarity or difference between a verb's two existing stems which ultimately determines the morphological fate of its infinitive in VMü.

Even if this proposal does have merit, given the paucity of our material it would be premature to attempt to identify what might constitute *sufficient* phonological similarity in this context, and I will not attempt to do so here. However, I close by observing that the notion of phonological similarity may also help to explain why *və'łajr* 'want' also retains its distinctive second conjugation infinitive ending, unlike e.g. *'tažər* 'be silent' < *tacēre* (Decurtins 1958: 132, 84). In the context of such forms, it is hard to see what would be 'wrong' with a potential **'və'łər* in its own right, and there is no inflectional augment to be identified in its paradigm. What does distinguish the paradigm of *və'łajr*, however, is the relative

irregularity of its other inflected forms, such as 3sg present *vo* (rather than e.g. *'vøʎ*; Decurtins 1958: 134). This may hold the key to its failure to adopt the innovative infinitive form.

If the infinitive form of verbs in VMü is genuinely determined at least in part by this intraparadigmatic factor of relative phonological similarity, this is a striking finding and one which, like the apparent trigger for the change in the first place, is unique in the context of the cases examined in this thesis. However this may be, the observation that the distribution of third conjugation infinitive morphology over the lexicon is sensitive to the presence of inflectional augments, as made separately by Eichenhofer 1999 and Maiden 2011b, has clearly merited re-examination. It emerges that the direct relevance of the Romance augment, as an inherited morphological entity of wide historical-comparative interest, is far from clear here; moreover, in the first conjugation cases we seem not to be dealing with ‘the inflectional augment’ in this standard sense at all, although in the behaviour of the verbs involved we may be witnessing evidence for the reinterpretation of pre-existing lexical material as a new augment behaving in parallel with the inherited one. More generally, we have also seen reason to doubt the idea that VMü shows particular conservatism in retaining traces of an earlier pan-Romansh state of affairs now submerged in the other varieties – a finding of interest both in the Romansh context and for Romance historical morphology more widely.

6. Conclusions

The present work has treated the phenomenon of heteroclisis from a number of different angles. Building on the interest accorded to unusual and ‘non-canonical’ morphological behaviour in general in the modern typological literature, it discusses both descriptively and theoretically oriented recent treatments of heteroclisis as a cross-linguistically observable feature of inflectional morphology comparable to other paradigmatic phenomena, alongside an investigation of different theoretical approaches to the more basic notion of inflectional class; considers the history of the term *heteroclisis* itself in grammatical thought and the appropriateness of its use in reference to different examples of inflectional behaviour; and, most substantially, provides several case studies investigating heteroclitite phenomena in the verb system of Latin and the Romance family in terms of their history and synchronic status.

This is the first treatment to explore in such detail the modern morphological conception of heteroclisis in its diachronic and synchronic aspects, as it can be identified as a recurrent phenomenon within a single domain of a particular language family over its recorded history. The potential scope of such a study is extremely broad, and naturally I have not attempted to discuss all the conjugational behaviour in the history of Latin and Romance varieties, or even just the languages treated here, to which the notion of heteroclisis could be applied. However, I have considered a range of cases which illustrate the various ways in which heteroclisis can arise in the course of a language’s development and the differing prominence that instances of heteroclitite behaviour can show in the context of closely related inflectional systems.

Moreover, as discussed in Chapter 2, while I took as my starting point for a presentation of the concept of heteroclisis the paradigms of Latin *domus* ‘house’ and Czech *pramen* ‘spring’, which inflect highly idiosyncratically for their respective

languages, I do not restrict my use of the term itself to such ‘unsystematic’, sporadic cases as does Noyer 2005. More importantly, I have not assumed that such phenomena are different in kind from those which affect more substantial groups of lexemes. This is partly a methodological issue, as my looser focus has allowed me to consider not only lexically unique or near-unique phenomena but also the behaviour of larger groups which play a more substantial role in the makeup of the inflectional system as a whole; but in fact we have seen that there may be cause to dispute the idea that any such distinction falls out naturally from the generative mechanisms of the grammatical system at large in the way Noyer suggests. Instead, I have drawn on the descriptive framework for inflectional morphology advanced by proponents of the Canonical Typology approach, according to which heterocclisis, like other inflectional phenomena such as suppletion and deponency, can be identified in terms of its relationship to ‘canonical inflection’, and thus propose that like these other phenomena, heterocclisis is instantiated by examples which can be identified as more or less highly canonical along particular dimensions: one of these dimensions does concern the number of lexemes sharing the relevant behaviour, but there are many more. For example, I have claimed that among examples of the Latin *petō* type treated in section 3.4, the relationship between Latin *quaerō* ‘seek’ and its perfect and third stem forms exemplified by *quaesīvī*, *quaesītum* (cf. Table 3.8 above) constitutes a less than perfectly canonical example of heterocclisis because the latter forms cannot be said to take all their inflectional properties from the regular behaviour of the fourth conjugation in these parts of the paradigm, meaning that a further type of stem alternation must also be identified here which is unmotivated by the presence of heterocclisis itself. Similarly, I have suggested that heterocclisis is more highly canonical when the inflectional classes implicated can be identified

unambiguously: this is the case e.g. for the *petō* type, but not for the *domō* type (section 3.3), as forms such as perfect *domuī* – while distinctively failing to pattern with the first conjugation – do not characterize any single other conjugation exclusively. Criteria such as these go some way towards setting out a synchronic canonical typology of heteroclisis, along the lines of those provided for suppletion and inflectional classes themselves by Corbett 2007a, 2009a. More generally, I have pointed out that Canonical Typology has not generally located heteroclisis in the right ‘place’ in the framework: as it depends on the idea of inflectional class, itself dependent on the non-canonical phenomenon of inflectional allomorphy, heteroclisis exists at a higher order of exceptionality than the other types of inflectional phenomena generally treated in terms of the canonical approach.

The broad historical scope of this work, with its attention to the diachrony and structural context of particular developments, has meant that the various case studies presented have privileged different kinds of evidence and discussion. Although this thesis is not intended primarily as a work of dialectology, the observations made in Chapter 4 on several patterns of heteroclite behaviour in the Romanian varieties depended on the presentation of (published and unpublished) dialect survey data precise to the level of individual villages, alongside the facts of the standard language and historical attestations from manuscripts dating back to the sixteenth century; Esher 2012 was able to draw on equally localized records in identifying heteroclisis in the course of her study of Occitan verb morphology, discussed in section 2.6.3. For Romansh (Chapter 5) we do not possess the same wealth of fine-grained dialectal data relevant to differences in verb morphology in particular, but we can rely on a wide range of pedagogical grammars of the (largely) standardized dialects representing the five dialect areas of the Romansh-speaking

region, along with some less systematic survey data and more general studies of local varieties, as well as historical evidence for verb morphology in earlier centuries. Of course, the treatment of developments in the verb system in these languages benefits from the fact that we also know a great deal about the parent system they all share as it is represented in our Latin attestations: the conjugational behaviour seen in Latin can be taken as the common ‘starting point’ for the various developments seen in Romance. On the other hand, to understand the background to this Latin conjugational system itself requires not only familiarity with attested developments in the early history of the language but also knowledge of the behaviour of the verb in proto-Indo-European and later in proto-Italic, for which no direct evidence can be provided, but which can be reconstructed to an extent thanks to the comparison of sister languages. The approach and argumentation applied to the cases featuring in Chapter 3 are thus quite different from those seen in the later chapters.

Along with these natural differences of focus, the case studies presented here have sometimes allowed for the discussion of highly particular questions whose resolution is principally relevant only within the linguistic context at issue, rather than to the behaviour of heteroclisis in general – while being of great interest within that context, for reasons which may go beyond any connection with heteroclisis. For example, as part of my treatment of heteroclite Latin verbs of the *domō* type, in section 3.3.2 I argued against the suggestion that a proposed deverbal suffix **-āye/o-* with atelic semantics is one of the chief sources of the Latin first conjugation, as de Vaan’s (2012) presentation of the case for this suffix fails to explain why the *domō* type should exist at all; I prefer an account which, while doing without this suffix, also gives a principled phonological reason for the existence of the *domō* type and much of its lexical makeup (Rix 1999). The specific arguments made cannot be

generalized beyond the immediate case, but deciding between the two positions is valuable in its own right, and the discussion is informed by a recognition of the heteroclite type's relevance to the question. Similarly, the debate surrounding the origins of the *capiō* type presented in section 3.2.3 is tightly bound up with questions about the plausibility of an inherited present stem formation making use of an athematic **-i-*, ostensibly reflected elsewhere in Indo-European: as I point out, following Martzloff 2006, the strikingly uniform root shape shown by the Latin verbs belonging to this type provides evidence in favour of a position which holds that no such formation in **-i-* existed, which if true has important consequences for the history of verb morphology in other Indo-European branches. Meanwhile, we have seen in section 5.4 that the attempt to characterize appropriately the conditions on the innovative lexical distribution of third conjugation infinitive morphology in the Romansh of Val Müstair touches on a much broader question concerning the history of the augment, one of the most characteristic features of Romance verb morphology.

This is merely to say that the findings from an investigation of any single instance of heteroclisis may naturally make their own specialized contribution alongside what they tell us about the more general phenomenon. This much was clear in the very first substantial presentation of a concrete case of heteroclite behaviour provided in this thesis, in section 2.6.1, where we saw that on the basis of the distribution of heteroclisis involving 'neuter *i*-stem' morphology over the Modern Greek dialects Karatsareas 2011a identifies a closer historical relationship between Asia Minor Greek and the Lesbian and Samian dialects than has generally been recognized. Inevitably, the treatment of any linguistic development will generate case-specific findings of this kind, and many such results have been

presented within their own context in the course of this thesis. At the same time, however, we have now seen enough examples of heteroclisis from Latin and Romance (and beyond), in the cases examined here as well as in those treated in previous work, to be able to bring together some more general observations concerning the ways in which heteroclisis may operate in diachrony and synchrony: how novel heteroclite patterns can originate, how their prevalence over the lexicon can alter, and how heteroclisis can interact with other characteristic features of a given inflectional system.

I begin by considering the range of factors which may lead to the emergence of heteroclisis, as identified chiefly from the history of the Latin and Romance verb systems. One important trigger, whose effects are seen in several of the cases treated here, is the operation of neutralizing sound change, which can affect the morphology by effacing what were previously clear-cut morphological distinctions. As noted already by Maiden 2009, and treated in section 4.2 above, the heteroclite behaviour widely observable in dialectal Romanian in the paradigms corresponding to standard *a coase* ‘sew’ and *a țese* ‘weave’ traces its origins ultimately to a historical process of vowel centralization affecting front vowels after *s*: heteroclisis arose because some of the phonological effects produced by this development in the third conjugation paradigms of these two verbs, notably the appearance of *-ă* rather than *-e* in the 3sg present, could be reinterpreted as morphological in the light of the existing distinctions found between the conjugation classes of Romanian; this innovative inflectional affiliation could then be generalized analogically to other parts of the paradigm. Following Zamfir (2005: 207) and others, I show that a similar explanation in terms of morphological ambiguity as the result of sound change is also likely to hold for the heteroclite behaviour of standard Romanian *a acoperi*

‘cover’ and other verbs of a similar shape, as discussed in section 4.4. Meanwhile, in section 5.4 I propose that in the Romansh of Val Müstair the prevalence of apparently third conjugation infinitives, in verbs otherwise showing the inflection proper to other conjugations, is also likely to stem from the morphologization of an originally phonological neutralization process of a slightly different kind, by which the final stress characteristic of infinitives outside the third conjugation was automatically retracted in certain prosodic contexts; once reinterpreted as a morphological phenomenon, this retraction was generalized to all contexts.

The potential role of sound change in the emergence of heteroclisis is also illustrated in the history of the Latin verb system. Section 3.2.2 has shown how the verbs of the *capiō* type, which synchronically draw on both third and fourth conjugation morphology, owe this heteroclitite status to the cluster of related sound changes, together termed ‘vowel weakening’, which affected vowels in non-initial syllables early in the attested history of the language; these changes automatically served to efface the formal distinctions between the inherited inflectional behaviour shown by these verbs (already partly shared with the productive fourth conjugation thanks to still earlier sound changes) and the behaviour seen in the diachronically unrelated third conjugation, with the result that most likely not a single inflectional form survived which could be identified as proper to the *capiō* type in its own right. Again, then, sound change gave rise to heteroclisis by eliminating an earlier conjugational distinction in certain cells. However, as is clear from these diachronic details, the role played by sound change in this Latin case is somewhat different from that seen in both Romanian and the Romansh of Val Müstair; the Latin development affected a set of verbs whose inflectional behaviour over the paradigm was already distinct from that shown by both the third and fourth conjugations before the sound

change took place, meaning that in giving rise to heteroclisis the neutralization here simply altered the morphological nature of this pre-existing distinction.

However, we have also seen evidence that the phonology of a lexeme can lead to its acquiring heteroclite inflection for reasons that do not involve sound change. Consider the phenomenon observed widely across the Occitan-speaking area by Esher 2012 and presented in section 2.6.3, whereby certain third conjugation verbs have taken on innovative conjugational behaviour in the future and conditional forms. It appears that heteroclisis emerged in these cases because of the relatively unfamiliar shape of the inherited future and conditional in the verbs affected; the theme vowel *e* expected in these forms is liable to be replaced by the much more lexically prevalent *i* or *a*, which are proper to different conjugations. Crucially, however, in the dialects where this heteroclisis is observed, the presence of the theme vowel *e* as opposed to zero in these forms is itself conditioned by stem shape. In this sense, the phonology of the lexeme is indirectly responsible for its development of a heteroclite paradigm. Meanwhile, the heteroclisis in nominal paradigms in dialectal Greek treated by Karatsareas 2011a seems to have originated in lexemes sharing a particular stress pattern, in which stress falls on the antepenultimate syllable in the nominative; as discussed in section 2.6.1, such nouns were susceptible to the shift in conjugational behaviour not because of sound change, but on the basis of developments internal to the history of Greek morphology, and the development of heteroclisis here appears to represent a morphological ‘repair strategy’ comparable to the rise of defectiveness in nominal paradigms elsewhere in Greek. In these two cases, heteroclisis thus emerged directly as a morphological phenomenon rather than from the reinterpretation of a change based in the phonology, but nonetheless targeted lexemes of a particular shape.

In some cases it is more difficult to identify the particular motivation behind the original emergence of heteroclite behaviour in a language: this is inevitably the case when we cannot even be certain where in the lexicon the behaviour originated. This is the case, for example, for the heteroclisis displayed by the Latin *petō* type, whose uncertain origins were examined in section 3.4. The ‘pan-Romansh’ phenomenon of section 5.2, by which fourth conjugation morphology has entered the paradigm of e.g. Vallader *cuorrer* < third conjugation *currere*, is also difficult to pin down to a particular starting point, and there is no guarantee even that it originally affected the same lexemes in the different dialects. Nonetheless, it seems unlikely that sound change is to blame, as there is no independent evidence that any appropriate sound change took place in the history of the dialects concerned; instead, I have suggested that the development was entirely morphological from the outset and reflected a general bias identifiable in the language towards the introduction of fourth conjugation morphology in place of that of the unproductive third (/second) conjugation; this can be stated as a move in the direction of system-dependent naturalness. Ultimately, however, the phenomenon may have arisen thanks to inherited lexical idiosyncrasies which characterize the morphology of individual items outside Romansh as well and for which no general Romansh rationale can be found.

At the same time, this thesis has uncovered individual instances of heteroclisis where the directly morphological introduction of innovative conjugational behaviour into the paradigm can indeed be securely ascribed to lexically specific causes. Three clear-cut cases of this kind feature in section 4.3, and involve the analogical influence of particular Romanian verbs on others originally belonging to different conjugations; our examples attest to the relevance of both

formal and semantic motivation in producing this influence. In Banat, we have seen (cf. Table 4.12) that the lexeme corresponding to standard Romanian first conjugation *a sughița* ‘hiccup’ frequently shows the incursion of fourth conjugation morphology into the paradigm, reflecting the influence of the more common verb *a înghiți* ‘swallow’: here both the related meanings and the similar stem shapes have promoted the transfer of inflectional behaviour from one lexeme to another. But shared semantics is enough in its own right to motivate this kind of influence, as shown by the existence of heteroclite behaviour in some varieties of Crișana in the lexeme *a zbură* with the meaning ‘jump’, which has been brought about thanks to the analogy of *a sări* ‘jump’ (cf. Table 4.16). Perhaps even more surprisingly, formal similarity alone can be sufficient: thus both *a mâna* ‘drive’ and *a amâna* ‘postpone’ show widespread heteroclisism simply because they coincidentally share their stem shape with the unrelated *a rămânea* ‘stay’ (cf. Table 4.17).

To sum up the foregoing: it appears that a distinction can be made between those instances of heteroclisism that arose directly as a result of morphological change, i.e. the substitution of one type of inflectional behaviour with another in part of the paradigm, and those that can ultimately be attributed to a purely phonological effect, which was susceptible to reinterpretation as morphological in nature in the light of existing features of the inflectional system. This can be compared with the diachronic distinction made with reference to suppletion, and noted in section 2.5, between those instances which come about thanks to the ‘incursion’ of outside forms into the inflectional paradigm and those which are due to sound change rather than any initial morphological disruption (Plank 1996, Juge 1999).

At the same time, neither of these categories should be treated as monolithic. As the examples treated here suggest, precisely because the potential for

morphological ambiguity depends on the contingent features of the phonology and the inflectional system in the first place, the phonological triggers for heteroclisis and the way in which they operate may have little in common from case to case, save that they will involve the neutralization of some morphologically relevant distinction in at least some contexts. Meanwhile, heteroclisis identifiable as a morphological effect from the outset can itself be attributed to numerous diverse causes, from the operation of lexically specific analogical influences (themselves motivated in various ways) to pressures from more general morphological trends or avoidance strategies. Thus one important aspect of heteroclisis as a cross-linguistic phenomenon to emerge from the present work is in fact the wealth of different ways in which it can be observed to arise even within the history of the verb in a few representatives of a single linguistic branch, and a more substantial survey within Latin/Romance or more widely would undoubtedly uncover new examples; in fact, I return below to another example of the emergence of heteroclisis which has not been treated here.

Moreover, the existence of a logical distinction between these two broad types should be set against the fact that, as we have seen, the detailed diachronic story behind the details of any given instance of heteroclisis may involve the interplay of phonological and morphological factors. This is perhaps best seen in the case of the Latin *domō* type treated in section 3.3. Here the immediate cause of the heteroclitite behaviour observed can be taken to be the effect of the historically attested, regular ‘weakening’ sound change which neutralized the distinction between *a* and *e* in a variety of phonological contexts – which happened to include those relevant to the *perfectum* and third stems of verbs such as *domō* ‘tame’, eventually giving *domuī*, *domitus* (<*domaw-, *domat-) morphologically indistinguishable from *monuī*, *monitus* (< *monew-, *monet-) to present *moneō*

‘warn’; this in turn was subsequent to a previous regular sound change which had collapsed the distinction in the present stem between the type **domaye/o-* and the ancestors of the productive first conjugation verbs in **-āye/o-*. Hence the small group of verbs originally displaying a series of three stems all containing short **a* came to retain no distinct formal characteristics of its own, instead ‘falling in’ with more widespread patterns in a way which divided the paradigm into two along the lines laid down by the pre-existing stem distribution. To this extent, the explanation is phonological: all the verbs originally involved here showed this **a*, which reflected their common feature of a laryngeal consonant root-finally at a still earlier stage. However, this small group itself had (at least) two sources with different morphological histories: in one, most likely exemplified by *domō*, the *perfectum* and third stems were themselves created productively within the history of Italic on the basis of a morphological reanalysis of the present stem; in the other, exemplified by *sonō* ‘sound’, the *perfectum* stem was inherited from the Indo-European parent language, and it served as the counterpart to two separate inherited present stems, both of which survive into our attestations. This explains the connection visible in Latin between heteroclisis of the *domō* type and the rare *Simultanperfekt* phenomenon. A full account of the emergence of this type of heteroclisis, and its role in the synchronic system of Latin, thus depends on the recognition of both distinctive phonological and morphological factors.

We saw above that in Romanian various instances can be found of heteroclisis emerging as the result of the direct morphological influence of one lexeme on another, where both previously inflected normally for their respective conjugations. However, it is also possible to detect the spread through the lexicon of heteroclisis as such, a process which again may operate at a lexically individual

level. The *domō* type appears to demonstrate this: I have suggested that *necō* ‘kill’, *fricō* ‘rub’, *plicō*, ‘fold’, which do not belong with the *domō* type etymologically, owe their heteroclite behaviour to the influence of *secō* ‘cut’ (and perhaps *micō* ‘glitter’) thanks to a combination of formal and semantic likeness. That heteroclisis has advanced through the lexicon in this way is clearer in the case of the *petō* type (section 3.4), where again form and semantics are implicated, potentially in combination, and it is likely that a single lexeme provided the original model for the heteroclite behaviour observed. Our studies of Romanian and especially Romansh, however, have shown that heteroclite behaviour can also spread through the lexicon in a less constrained way, to the extent that in Surselvan an originally heteroclite pattern has come to represent the new norm, apparently not by means of a ‘single’ principled change (as has perhaps been the case in Val Müstair) but as the result of a gradual spread through the lexicon whose earlier stages are visible in Vallader and which is still not complete (though the word is perhaps inappropriately teleological) owing to the relatively conservative behaviour of a few Surselvan items such as *metter* ‘put’. These Surselvan and VMü cases instantiate an important point, and one which I have attempted to draw out at various points in the present work: while it is reasonable to consider the kind of partial inflectional class changes treated here as the diachronic counterpart to the synchronic phenomenon of heteroclisis, the two are logically and empirically distinct.

As well as the distribution of heteroclite behaviour over the lexicon, however, we have also been able to investigate the distribution of innovative conjugational morphology over the *paradigm* of the lexemes affected. Especially notable are those findings from Romanian, taking up much of chapter 4 in the present work, in which – in several different verbs of different shapes – etymologically aberrant

conjugational affiliation can be seen to have originated in certain parts of the paradigm for phonological reasons, but has subsequently spread to new cells, resulting in patterns of conjugational alternation for which sound change cannot be directly responsible. To observe this phenomenon in a given instance is of linguistic value in itself, as it demonstrates that speakers have reinterpreted the original development in question as a morphological one, although this need not have been the case. But, following Maiden 2009, I have drawn attention in particular to the striking agreement shown by the Romanian varieties investigated, both dialectal and standard, in their redistribution of conjugational affiliation over the paradigm in these cases. This most commonly implicates a set of cells – the so-called ‘N-pattern’ cells introduced in section 2.5 above – which do not form a natural class in morphosyntactic terms, but are nonetheless found to show coherent morphological behaviour both within Romanian and more widely across Romance, for example with regard to stem selection even in cases of incursive suppletion; the widespread realignment of heteroclite behaviour, originally arising on phonological grounds, into innovative conformance with this pattern is an indication that it is not merely a contingent, fossilized feature of the language but continues to play an organizational role in speakers’ conception of the inflectional system. Notably, the pan-Romansh heteroclisism treated in section 5.2, while different in other respects from that seen in Romanian – it is unlikely to have arisen for phonological reasons and involves the spread of fourth rather than first conjugation morphology – is sensitive to the same morphological division within the paradigm. Within Romanian, I have identified new evidence from the Banat varieties (section 4.3) also pointing to the importance of this structural zone, here on the basis of heteroclisism arising the direct influence of one

lexeme (fourth conjugation *a inghiṭi* ‘swallow’) on another (first conjugation *a sughiṭa* ‘hiccup’).

The observation that the intraparadigmatic splits involved in instances of heteroclisis may not only divide lexemes along morphosyntactic lines, as is arguably the case for the Latin *domō* and *petō* types, but also along purely morphological lines is an important one; it disfavours the conception of heteroclisis as licensed by Carstairs-McCarthy’s ‘slabs’, in favour of an approach which promotes the existence of reliable implicational relationships between cells. Such morphomic patterns can be recognized as *natural* features of a given language in the sense that they represent system-dependent structural properties in the terms of Natural Morphology (section 2.2.2): in finding that heteroclisis falls into patterns which accord with what is structure-congruous given the existing properties of the system, we have observed the relevance of system-specific naturalness, as is already visible in the structurally coherent behaviour of other inflectional features which have been recognized in the literature as displaying morphomic patterning over the paradigm, such as suppletion (section 2.5).

In the broader perspective, then, the present findings help to cement the position, adopted by e.g. Corbett 2015, that heteroclisis can usefully be ranged alongside other inflectional characteristics which entail ‘lexical splits’ in the paradigm, and that it can be sensitive to the same diachronic pressures and take on the same synchronic distribution as they can, pointing to the productivity and psychological reality of certain structural aspects of the morphological system. More narrowly, the foregoing studies have helped to shed light on individual questions of Latin and Romance historical linguistics. Naturally, however, both specific and general questions remain. Within the history of Latin and Romance in particular, it

would be particularly interesting to investigate from the point of view of heteroclisis a phenomenon which has been touched on several times in the course of the present work, namely the rise of the ‘inflectional augments’ seen in Romanian and Romansh. These, and their Latin origins, have featured here as part of a discussion of the basic features of conjugation in those languages, but also with regard to the specific conditioning on the behaviour of the infinitive in the Romansh dialect of Val Müstair. However, one point it has not been possible to address in the present work is that the early history of the augment in *-ēs̄c/īs̄c-* can itself be observed to involve diachronic conjugational class change: verbs bearing this element in Latin belonged to the *third* conjugation, and yet came to form part of productive fourth conjugation morphology in many Romance languages.¹⁵⁴ This implies a historical stage of heteroclisis, and one hardly known to us from written attestations, where innovative paradigms were emerging from the combination of pairs of lexemes which belonged to different conjugations within the synchronic system, illustrating a different origin for heteroclisis than any of those treated in the present work. While the story of the Romance augments has recently received a thorough treatment in Meul 2013, the possibility is worth pursuing further that a survey of the behaviour of the fourth conjugation augment across Romance would uncover synchronic traces of these heteroclitite origins.

Meanwhile, beyond the fact that a study such as this can only cover a sample of the potential material available to it, I have systematically selected instances of heteroclisis which play, or diachronically come to play, a reasonably prominent role in the inflectional system as a whole, i.e. which are *less* lexically exceptional and thus less canonical according to that criterion. It would be desirable to investigate

¹⁵⁴ Moreover, this fourth conjugation marker can be seen to have spread in turn to the first conjugation in Romansh (cf. section 5.2); again, historical attestations of the course of this change are lacking.

more highly exceptional cases in their own right to determine whether the apparent lack of any obvious motivation, even morphology-internal, which can be identified in cases like that of Latin *domus* is in fact representative of such cases, and whether the historical factors which give rise to them are systematically different from those which characterize better-populated types. A closely connected question is whether there are particular types of paradigmatic patterns in heteroclisis (or indeed other types of lexical split) that can arise without being liable to be reshaped by the morphology into productive patterns and thereby made 'available' to the wider lexicon, as was seen in Romanian. The treatment of such questions requires the survey of a wider range of material than is provided by the inflectional history of the Latin/Romance verb: I hope to have demonstrated, however, that even the investigation of this single family from the point of view of heteroclisis in diachrony and synchrony provides much that is of interest for the study of inflectional morphology.

Bibliography

- Ackerman, Farrell, Blevins, James P., and Malouf, Robert (2009). 'Parts and wholes: implicative patterns in inflectional paradigms', in J. P. Blevins and J. Blevins (eds.), *Analogy in grammar: form and acquisition*. Oxford: Oxford University Press, 54-82.
- Ackerman, Farrell and Malouf, Robert (2013). 'Morphological organization: the low conditional entropy conjecture.' *Language* 89(3): 429-64.
- Ackerman, Farrell and Malouf, Robert (2015). 'The No Blur Principle effects as an emergent property of language systems', in A. E. Jurgensen, H. Sande, S. Lamoureux, K. Baclawski, and A. Zerbe (eds.), *Proceedings of the forty-first annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 1-14.
- Adams, James N. (2007). *The regional diversification of Latin, 200 BC-AD 600*. Cambridge: Cambridge University Press.
- Adams, James N. (2013). *Social variation and the Latin language*. Cambridge: Cambridge University Press.
- Allanche, Louis (1941). *Éléments de grammaire du dialecte quercynol des environs de Moissac et de Montauban suivis de proverbes et de dictons*. Montauban: Forestié.
- Alibèrt, Loís (1976). *Gramatica occitana segon los parlars lengadocians*. 2nd edition. Montpellier: CEO.
- Allen, Joseph H. and Greenough, James B. (1903). *New Latin grammar* [revised by J. B. Greenough, G. L. Kittredge, A. A. Howard, and B. L. D'Ooge.] Boston: Ginn and Co.
- ALLOc* = Ravier (1978-93).
- ALLOr* = Boisgontier (1981-86).
- Anderson, Stephen R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. (2008). 'Phonologically conditioned allomorphy in the morphology of Surmiran (Rumantsch).' *Word Structure* 1(2): 109-34.
- Anderson, Stephen R. (2011). 'Stress-conditioned allomorphy in Surmiran (Rumantsch)', in M. Maiden et al. (eds.), 13-35.
- Anderson, Stephen R. (2013). 'Stem alternations in Swiss Rumantsch', in S. Cruschina et al. (eds.), 8-23.

- Aronoff, Mark (1994). *Morphology by itself: stems and inflectional classes*. Cambridge, Mass.: MIT Press.
- Arquint, Jachen C. (1964). *Vierv ladin: grammatica elementara dal rumantsch d'Engiadina bassa*. Tusan: Lia Rumantscha.
- Ascoli, Graziadio I. (1873). *Saggi ladini (Archivio glottologico italiano 1)*. Turin: Ermanno Loescher.
- Ascoli, Graziadio I. (1882-85). 'L'Italia dialettale.' *Archivio glottologico italiano* 8: 98-128.
- Asher, Ronald E. and Simpson, James M. Y. (eds.) (1994). *Encyclopedia of language and linguistics* [10 vols]. Oxford: Pergamon.
- Aski, Janice (1995). 'Verbal suppletion: an analysis of Italian, French and Spanish *to go*.' *Linguistics* 33(3): 403-32.
- Avram, Mioara (2001). *Gramatica pentru toți*. Bucharest: Humanitas.
- Baerman, Matthew (2007). 'Morphological typology of deponency', in M. Baerman et al. (eds.), 1-19.
- Baerman, Matthew (2012). 'Paradigmatic chaos in Nuer.' *Language* 88(3): 467-94.
- Baerman, Matthew (2014). 'Covert systematicity in a distributionally complex system.' *Journal of Linguistics* 50(1): 1-47.
- Baerman, Matthew, Brown, Dunstan, and Corbett, Greville G. (2005). *The syntax-morphology interface: a study of syncretism*. Cambridge: Cambridge University Press.
- Baerman, Matthew, Corbett, Greville G., Brown, Dunstan, and Hippisley, Andrew (eds.) (2007). *Deponency and morphological mismatches*. Oxford: Oxford University Press.
- Baerman, Matthew, Corbett, Greville G., and Brown, Dunstan (eds.) (2010). *Defective paradigms: missing forms and what they tell us*. Oxford: Oxford University Press.
- Bailey, Charles-James N. (1980). 'The role of language development in a theory of language.' *Papere zur Linguistik* 22(1): 33-46.
- Bailyn, John and Nevins, Andrew (2008). 'Russian genitive plurals are impostors', in A. Bachrach and A. Nevins (eds.), *Inflectional identity*. Oxford: Oxford University Press, 237-70.
- Bakker, Pieter (2003). 'Mixed languages as autonomous systems', in Y. Matras and P. Bakker (eds.), *The mixed language debate*. Berlin: Mouton de Gruyter, 107-50.

- Beekes, Robert S. P. (2011). *Comparative Indo-European linguistics: an introduction*. 2nd edition. Amsterdam: John Benjamins.
- Beltechi, Eugen, Faiciuc, Ioan, Mocanu, Nicolae, and Neiescu, Petru (eds.) (2005). *Noul atlas lingvistic român pe regiuni. Banat*, vol. 4. Bucharest: Editura Academiei Române.
- Benoît, Robert and Lavaud, René (1932). *Abrégé de grammaire périgourdine*. Périgueux: Éditions du Bournat du Périgord.
- Benveniste, Émile (1935). *Origines de la formation des noms en indo-européen*. Paris: Adrien-Maisonneuve.
- Bertacca, Antonio (2009). *Natural Morphology and the loss of nominal inflections in English*. Pisa: PLUS - Pisa University Press.
- Birkmann, Thomas (1987). *Präteritopresentia: morphologische Entwicklungen einer Sonderklasse in den altgermanischen Sprachen*. Tübingen: Niemeyer.
- Bittner, Andreas (1985). 'Implikative Hierarchien in der Morphologie: das 'Stark-Schwach-Kontinuum' der neuhochdeutschen Verben.' *Acta Linguistica Academiae Scientiarum Hungaricae* 35(1/2): 31-42.
- Bittner, Andreas (1996). *Starke 'schwache' Verben - schwache 'starke' Verben: deutsche Verbflexion und Natürlichkeit*. Tübingen: Stauffenberg.
- Blaylock, Curtis (1975). 'The Romance development of the Latin verbal augment -sk-.' *Romance Philology* 28(4): 434-44.
- Blevins, James P. (2004). 'Inflectional classes and economy', in G. Müller, L. Gunkel, and G. Zifonun (eds.), *Explorations in nominal inflection*. Berlin: Mouton de Gruyter, 41-85.
- Bobaljik, Jonathan D. (2002). 'Syncretism without paradigms: Remarks on Williams 1981, 1994', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2001*. Dordrecht: Kluwer, 53-86.
- Bock, Bettina (2008). *Die einfach thematischen Präsentien in der dritten Konjugation des Lateinischen*. Graz: Leykam.
- Boisgontier, Jacques (1981-86). *Atlas linguistique et ethnographique du Languedoc oriental* [3 vols]. Paris: Éditions du CNRS.
- Bonami, Olivier and Boyé, Gilles (2010). 'Opaque paradigms, transparent forms: Nepali conjugation.' Paper presented at the *Fifth Workshop on Theoretical Morphology*, Lutherstadt Wittenberg, 26th June 2010.

- Booij, Geert E., Lehmann, Christian, Mugdan, Joachim, and Skopeteas, Stavros (eds.) (2004). *Morphologie: ein internationales Handbuch zur Flexion und Wortbildung / Morphology: an international handbook on inflection and word-formation* [2 vols]. Berlin: Mouton de Gruyter.
- Boretzky, Norbert and Iglá, Birgit (1994). 'Romani mixed dialects', in P. Bakker and M. Mous (eds.), *Mixed languages. 15 case studies in language intertwining*. Amsterdam: IFOTT, 35-68.
- Borodina, Melitina A. (1969). *Современный литературный ретороманский язык Швейцарии* [Sovremennyj literaturnyj retoromanskij jazyk Švejcarii]. Leningrad: Nauka.
- Brown, Dunstan and Chumakina, Marina (2012). 'What there might be and what there is: an introduction to Canonical Typology', in D. Brown, M. Chumakina, and G. G. Corbett (eds.), *Canonical morphology and syntax*. Oxford: Oxford University Press, 1-19.
- Brown, Dunstan and Hippišley, Andrew (2012). *Network morphology: a defaults-based theory of word structure*. Cambridge: Cambridge University Press.
- Brown, Edward K. and Anderson, Anne (eds.) (2006). *Encyclopedia of language and linguistics* [14 vols]. 2nd edition. Amsterdam: Elsevier.
- Buck, Carl D. (1904). *A grammar of Oscan and Umbrian: with a collection of inscriptions and a glossary*. Boston/London: Ginn.
- Candrian, J. P. (1900). *Der Dialekt von Bivio-Stalla*. Zürich: Halle.
- Caragiu-Marioțeanu, Matilda (1975). *Compendiu de dialectologie română*. Bucharest: Editura Științifică și Enciclopedică.
- Carstairs[-McCarthy], Andrew (1983). 'The paradigm economy principle.' *Journal of Linguistics* 19(1): 115-28.
- Carstairs[-McCarthy], Andrew (1984). 'Paradigm economy in the Latin third declension.' *Transactions of the Philological Society* 82(1): 117-37.
- Carstairs[-McCarthy], Andrew (1987). *Allomorphy in inflexion*. London: Croom Helm.
- Carstairs[-McCarthy], Andrew (1988a). 'Paradigm economy: a reply to Nyman.' *Journal of Linguistics* 24(2): 489-99.
- Carstairs[-McCarthy], Andrew (1988b). 'Some implications of phonologically conditioned suppletion', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 1988*. Dordrecht: Foris, 67-94.

- Carstairs-McCarthy, Andrew (1991). 'Inflection classes: two questions with one answer', in F. Plank (ed.), *Paradigms: the economy of inflection*. Berlin/New York: Mouton de Gruyter, 213-53.
- Carstairs-McCarthy, Andrew (1994). 'Inflection classes, gender and the Principle of Contrast.' *Language* 70(4): 737-88.
- Carstairs-McCarthy, Andrew (1998). 'How lexical semantics constrains inflectional allomorphy', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 1997*. Dordrecht: Kluwer, 1-24.
- Carstairs-McCarthy, Andrew (2010). *The evolution of morphology*. Oxford: Oxford University Press.
- Carstairs-McCarthy, Andrew (2014). 'Affixes and stem alternants in Latvian nouns: implications for inflectional theory.' *Baltic Linguistics* 5: 59-80.
- CIL I² = Königlich Preußische Akademie der Wissenschaften zu Berlin, *Corpus Inscriptionum Latinarum* (Berlin 1893-), vol. I².
- Clackson, James (2007). *Indo-European linguistics: an introduction*. Cambridge: Cambridge University Press.
- Clark, Eve E. (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.
- Collinge, Neville E. (1985). *The laws of Indo-European*. Amsterdam: John Benjamins.
- Comrie, Bernard (2001). 'How independent is Tsezic morphology?', in M. Andronis, C. Ball, H. Elston, and S. Neuvel (eds.), *CLS 37: the panels. Papers from the parasessions of the thirty-seventh meeting of the Chicago Linguistic Society*, vol. 2. Chicago: Chicago Linguistic Society, 367-83.
- Comrie, Bernard (2003). 'When agreement gets trigger-happy', in D. Brown, G. G. Corbett, and C. Tiberius (eds.), *Agreement: a typological perspective* (special issue of *Transactions of the Philological Society* 101(2)). Oxford: Blackwell, 313-37.
- Conrad, Rudi (1985). *Lexikon sprachwissenschaftlicher Termini*. Leipzig: Bibliographisches Institut.
- Corbett, Greville G. (2003). 'Agreement: canonical instances and the extent of the phenomenon', in G. E. Booij, J. DeCesaris, A. Ralli, and S. Scalise (eds.), *Topics in morphology: selected papers from the third Mediterranean Morphology Meeting (Barcelona, September 20-22, 2001)*. Barcelona: Universitat Pompeu Fabra, 109-28.
- Corbett, Greville G. (2007a). 'Canonical typology, suppletion and possible words.' *Language* 83(1): 8-42.

- Corbett, Greville G. (2007b). 'Deponency, syncretism, and what lies between', in M. Baerman et al. (eds.), 21-43.
- Corbett, Greville G. (2009a). 'Canonical inflectional classes', in F. Montermini, G. Boyé, and J. Tseng (eds.), *Selected proceedings of the sixth Décembrettes: morphology in Bordeaux*. Somerville, Mass.: Cascadilla, 1-11.
- Corbett, Greville G. (2009b). 'Suppletion: typology, markedness, complexity', in P. O. Steinkrüger and M. Krifka (eds.), *On Inflection*. Berlin: Mouton de Gruyter, 25-40.
- Corbett, Greville G. (2015). 'Morphosyntactic complexity: a typology of lexical splits.' *Language* 91(1): 145-93.
- Coteanu, Ion, Mareaș, Lucreția, and Avram, Larisa (1998). *Dicționar explicativ al limbii române*. 2nd edition. Bucharest: Univers Enciclopedic.
- Cowgill, Warren (1970). 'Italic and Celtic superlatives and the dialects of Indo-European', in G. Cardona, H. M. Hoenigswald, and A. Senn (eds.), *Indo-European and Indo-Europeans: papers presented at the third Indo-European Conference at the University of Philadelphia*. Philadelphia: University of Pennsylvania Press, 113-53.
- Cruschina, Silvio, Maiden, Martin, and Smith, John Charles (eds.) (2013). *The boundaries of pure morphology: diachronic and synchronic perspectives*. Oxford: Oxford University Press.
- Dalbera, Jean-Philippe (1994). *Les parlers des Alpes-Maritimes : étude comparative, essai de reconstruction*. London: AIEO.
- Dammel, Antje (2011). *Konjugationsklassenwandel: Prinzipien des Ab-, Um- und Ausbaus verbalflexivischer Allomorphie in germanischen Sprachen*. Berlin: De Gruyter.
- Dammel, Antje, Nowak, Jessica, and Schmuck, Mirjam (2010). 'Strong verb paradigm leveling in four Germanic languages. A category frequency approach.' *Journal of Germanic Linguistics* 22(4): 341-65.
- Decurtins, Alexi (1958). *Zur Morphologie der unregelmäßigen Verben im Bündnerromanischen: historisch-deskriptive Studie mit besonderer Berücksichtigung des Sur- und Sutselvischen*. Bern: Francke.
- Denooz, Joseph (2010). *Nouveau lexique fréquentiel du latin*. Hildesheim: Olms.
- de Vaan, Michiel A. C. (2008). *Etymological dictionary of Latin and the other Italic languages*. Leiden: Brill.
- de Vaan, Michiel A. C. (2011). 'PIE *i*-presents, *s*-presents, and their reflexes in Latin.' *Glotta* 87(1): 23-36.

- de Vaan, Michiel A. C. (2012). 'Latin deverbal presents in *-ā-*', in H. C. Melchert (ed.), *The Indo-European verb. Proceedings of the conference of the Society for Indo-European Studies, Los Angeles, 13-15 September 2010*. Wiesbaden: Reichert, 315-32.
- Diaconescu, Ion (1977). *Infinitivul în limba română*. Bucharest: Editura Științifică și Enciclopedică.
- Diaconescu, Rodica C. (1999). *Romanian imperatives from the perspective of generative grammar*. MA thesis, University of Ottawa.
- Donegan, Patricia and Stampe, David (1979). 'The study of Natural Phonology', in D. Dinnsen (ed.), *Current approaches to phonological theory*. Bloomington: Indiana University Press, 126-73.
- Donohue, Mark (2001). 'Animacy, class and gender in Burmeso', in A. Pawley, M. Ross, and D. Tryon (eds.), *The boy from Bundaberg: studies in Melanesian linguistics in honour of Tom Dutton*. Canberra: Pacific Linguistics, 97-115.
- Dressler, Wolfgang U. (1985a). *Morphology: the dynamics of derivation*. Ann Arbor: Karoma.
- Dressler, Wolfgang U. (1985b). 'Typological aspects of Natural Morphology.' *Acta Linguistica Academiae Scientiarum Hungaricae* 35(1/2): 51-70.
- Dressler, Wolfgang U. (2002). 'Latin inflection classes', in A. M. Bolkestein, C. H. M. Kroon, H. Pinkster, H. W. Remmelink, and R. Risselada (eds.), *Theory and description in Latin linguistics*. Amsterdam: Gieben, 91-110.
- Dressler, Wolfgang U. (2003). 'A sketch of dynamic morphology of German verb inflection', in H. G. Jacobsen, D. Bleses, T. O. Madsen, and P. Thomsen (eds.), *Take Danish – for instance: linguistic studies in honour of H. Basbøll presented on the occasion of his 60th birthday, 12 July 2003*. Odense: University Press of Southern Denmark, 29-39.
- Dressler, Wolfgang U., Kilani-Schoch, Marianne, Gagarina, Natalia, Pestal, Lina, and Pöchtrager, Markus (2006). 'On the typology of inflection class systems.' *Folia Linguistica* 40(1/2): 51-74.
- DRG* = von Planta and Melcher (1939-).
- Eichenhofer, Wolfgang (1999). *Historische Lautlehre des Bündnerromanischen*. Tübingen: Francke.
- Eichner, Heiner (1974). *Untersuchungen zur hethitischen Deklination*. Doctoral thesis, Universität Erlangen – Nürnberg.
- Embick, David and Noyer, Rolf (2007). 'Distributed Morphology and the syntax/morphology interface', in G. Ramchand and C. Reiss (eds.), *The Oxford handbook of linguistic interfaces*. Oxford: Blackwell, 289-324.

- Emonds, Joseph E. and Spaelti, Philip (2005). 'Fully distributing morphology: the phonology and syntax of Latin case inflections.' *Theoretical and applied linguistics at Kobe Shoin* 8: 1-20.
- Enger, Hans-Olav (1998). *The classification of strong verbs in Norwegian with special reference to the Oslo dialect. A study in inflectional morphology*. Oslo: Scandinavian University Press.
- Enger, Hans-Olav (2007). 'The No Blur Principle meets Norwegian dialects.' *Studia Linguistica* 61(3): 278-309.
- Enger, Hans-Olav (2013). 'Vocabular clarity meets Faroese noun declensions.' *Folia Linguistica* 47(2): 345-74.
- Ernout, Alfred (1918). 'Cas en -e- et cas en -i- de la troisième déclinaison dans Lucrèce.' *Revue de philologie* 42(3): 133-68.
- Esher, Louise (2012). *Future, conditional, and autonomous morphology in Occitan*. Doctoral thesis, University of Oxford.
- Exon, Charles (1906). 'Medial vowel-syncope in Latin.' *Hermathena* 32: 117-43.
- Feldstein, Ronald (2004). 'On the structure of syncretism in Romanian conjugation', in J. Auger, J. C. Clements, and B. Vance (eds.), *Contemporary approaches to Romance linguistics*. Amsterdam: John Benjamins, 177-95.
- Finkel, Raphael A. and Stump, Gregory T. (2009). 'Principal parts and degrees of paradigmatic transparency', in J. P. Blevins and J. Blevins (eds.), *Analogy in grammar: form and acquisition*. Oxford: Oxford University Press, 13-53.
- Flobert, Pierre (1975). *Les verbes déponents latins des origines à Charlemagne*. Paris: Société d'Édition 'Les Belles Lettres'.
- Fortson, Benjamin W. IV (2004). *Indo-European language and culture: an introduction*. Oxford: Blackwell.
- Frawley, William (ed.) (2003). *International encyclopedia of linguistics* [4 vols]. 2nd edition. Oxford: Oxford University Press.
- Gabka, Kurt (1975). *Die russische Sprache der Gegenwart, vol. 2: Morphologie*. Düsseldorf: Brücken-Verlag.
- Ganzoni, Gian Paul (1977). *Grammatica ladina : grammatica sistematica dal rumauntsch d'Engiadin'Ota per scholars e creschieus da lingua rumauntscha e tudas-cha*. Chur: Lia Rumauntscha.
- Garde, Paul (1980). *Grammaire russe*. Paris: Institut d'études slaves.
- Giacalone Ramat, Anna and Ramat, Paolo (1998). *The Indo-European languages*. London: Routledge.

- Gönczöl-Davies, Ramona (2007). *Romanian: an essential grammar*. London: Routledge.
- Graur, Alexandre (1939). 'La IV. conjugaison latine.' *Bulletin de la Société de linguistique de Paris* 40: 127-150.
- Gregor, Douglas B. (1982). *Romontsch: language and literature: the Sursilvan Raeto-Romance of Switzerland*. Cambridge/New York: Oleander Press.
- Haiman, John and Benincà, Paola (1992). *The Rhaeto-Romance languages*. London: Routledge.
- Halle, Morris and Marantz, Alec (1993). 'Distributed Morphology and the pieces of inflection', in K. Hale and S. J. Keyser (eds.), *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press, 111-76.
- Halle, Morris and Marantz, Alec (1994). 'Some key features of Distributed Morphology', in A. Carnie and H. Harley (eds.), *MITWPL 21: papers on phonology and morphology*, Cambridge, Mass.: MITWPL, 275-88.
- Halle, Morris and Marantz, Alec (2008). 'Clarifying "blur": paradigms, defaults, and inflectional classes', in A. Bachrach and A. Nevins (eds.), *Inflectional identity*. Oxford: Oxford University Press, 55-72.
- Halle, Morris and Vaux, Bert (1998). 'Theoretical aspects of Indo-European nominal morphology: the nominal declensions of Latin and Armenian', in J. Jasanoff, H. C. Melchert, and L. Oliver (eds.), *Mir Curad: studies in honor of Calvert Watkins*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft, 223-40.
- Haspelmath, Martin (2002). *Understanding morphology*. London: Arnold.
- Heim, Michael H. (1982). *Contemporary Czech*. Columbus, Ohio: Slavica.
- Hinzelin, Marc-Olivier and Gaglia, Sascha (2012). 'Morphological theories, the autonomy of morphology, and Romance data', in S. Gaglia and M.-O. Hinzelin (eds.), *Inflection and word formation in Romance languages*. Amsterdam: John Benjamins.
- Huonder, Josef (1901). *Der Vokalismus der Mundart von Disentis*. Erlangen: Junge.
- HWR = Bernardi, Rut and Stricker, Hans (1994). *Handwörterbuch des Rätoromanischen: Wortschatz aller Schriftsprachen, einschliesslich Rumantsch Grischun, mit Angaben zur Verbreitung und Herkunft* [3 vols]. Zürich: Offizin.
- Hyman, Larry (2009). 'How (not) to do phonological typology: the case of pitch-accent.' *Language Sciences* 31(2/3): 213-38.
- Iliescu, Maria and Mourin, Louis (1991). *Typologie de la morphologie verbale romane*. Innsbruck: Amoe.

- Iordan, Iorgu (1935). 'Forme de conjugare mixtă în limba română.' *Buletinul institutului de filologie română "Alexandru Philippide"* 2: 47-127.
- Iordan, Iorgu (1948). *Limba română actuală : o gramatică a "greșelilor"*. 2nd edition. Iași: Editura Socec.
- Isebaert, Lambert (1994[1995]). 'Zur Stammbildung von lateinisch (°)*crepāre*.' *Münchener Studien zur Sprachwissenschaft* 55: 41-44.
- Jackendoff, Ray (2002). *Foundations of language*. Oxford: Oxford University Press.
- Jakobi, Angelika (1990). *A Fur grammar: phonology, morphophonology, and morphology*. Hamburg: Buske.
- Jasanoff, Jay H. (1978). *Stative and middle in Indo-European*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Jasanoff, Jay H. (2003). *Hittite and the Indo-European verb*. Oxford: Oxford University Press.
- Jasanoff, Jay H. (2008). 'The accentual type **vèdō*, **vedetì* and the origin of mobility in the Balto-Slavic verb.' *Baltistica* 43(3): 339-79.
- Juge, Matthew L. (1999). 'On the rise of suppletion in verbal paradigms', in S. S. Chang, L. Liaw, and J. Ruppenhofer (eds.), *Proceedings of the twenty-fifth annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 183-94.
- Juilland, Alphonse G. and Edwards, Prior M. H. (1971). *The Rumanian verb system*. The Hague/Paris: Mouton.
- Kammenhuber, Annelies (1955). 'Zur Genese der hethitischen -*r/-n*-Heteroclitica', in *Corolla linguistica: Festschrift Ferdinand Sommer zum 80. Geburtstag am 4. Mai 1955, dargebracht von Freunden, Schülern und Kollegen*. Wiesbaden: Harrassowitz, 97-106.
- Karatsareas, Petros (2011a). 'Neuter heterocclisis in Asia Minor Greek: origin and development.' *Νεοελληνική διαλεκτολογία* [Neoellinikí dialektología] 6: 111-35.
- Karatsareas, Petros (2011b). *A study of Cappadocian Greek nominal morphology from a diachronic and dialectological perspective*. Doctoral thesis, University of Cambridge.
- Katamba, Francis (2004). *Morphology*. London: Routledge.
- Kiparsky, Paul (2005). 'Blocking and periphrasis in inflectional paradigms', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Springer, 113-35.

- Kirby, Simon and Hurford, James R. (2002). 'The emergence of linguistic structure: an overview of the iterated learning model', in A. Cangelosi and D. Parisi (eds.), *Simulating the evolution of language*. London: Springer, 121-48.
- Kortlandt, Frederik (1989). 'Lithuanian *statyti* and related formations.' *Baltistica* 25(2): 104-12.
- Koziol, Heinrich (1872). *Der Stil des L. Apuleius. Ein Beitrag zur Kenntniss des sogenannten afrikanischen Lateins*. Vienna: Gerold.
- Kuryłowicz, Jerzy (1927). 'Les effets du *a* en indo-iranien.' *Prace Filologiczne* 11: 201-43.
- Laurent, Richard (1999). *Past participles from Latin to Romance*. Berkeley: University of California Press.
- Lausberg, Heinrich (1972). *Romanische Sprachwissenschaft*, vol. 3: *Formenlehre*. 2nd edition. Berlin: De Gruyter.
- Ledgeway, Adam (2012). *From Latin to Romance: morphosyntactic typology and change*. Oxford: Oxford University Press.
- Lehmann, Christian and Moravcsik, Edith (2000). 'Nouns', in G. E. Booij et al. (eds.), vol. 1, 732-57.
- Leumann, Manu (1977). *Lateinische Laut- und Formenlehre*. 5th edition. Munich: C. H. Beck.
- Lewis, Geoffrey (2000). *Turkish grammar*. 2nd edition. Oxford: Oxford University Press.
- Lieberman, Mark and Prince, Alan (1977). 'On stress and linguistic rhythm.' *Linguistic Inquiry* 8(2): 249-336.
- Lignières, Marcel (1951). *Grammaire pratique de langue d'oc, avec vocabulaire*. Béziers: Sudiep.
- LIV*² = Rix (2001).
- Liver, Ricarda (1991). *Manuel pratique de romanche : sursilvan – vallader*. 2nd edition. Chur: Lia Rumantscha.
- Liver, Ricarda (1999). *Rätoromanisch : eine Einführung in das Bündnerromanische*. Tübingen: Narr.
- Lombard, Alf (1954). *Le verbe roumain : étude morphologique*, vol. 1. Lund: Gleerup.
- Lombard, Alf (1955). *Le verbe roumain : étude morphologique*, vol. 2. Lund: Gleerup.

- Lombard, Alf (1974). *La langue roumaine. Une présentation*. Paris: Klincksieck.
- Loporcaro, Michele (2011). 'Phonological processes', in M. Maiden, J. C. Smith, and A. Ledgeway (eds.), *The Cambridge history of the Romance languages*, vol. 1: *Structures*. Cambridge/New York: Cambridge University Press, 109-54.
- Luis, Ana (2008). 'The effect of language contact on Romance verbal paradigms: an empirical survey.' *Linguística: revista de estudos linguísticos da Universidade do Porto* 3: 137-54.
- Maiden, Martin (1995). *A linguistic history of Italian*. London: Longman.
- Maiden, Martin (2001). 'A strange affinity: *perfecto y tiempos afines*.' *Bulletin of Hispanic Studies* 78(4): 441-64.
- Maiden, Martin (2004). 'Verb augments and meaninglessness in Romance morphology.' *Studi di grammatica italiana* 22: 1-61.
- Maiden, Martin (2005). 'Morphological autonomy in diachrony', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Springer, 137-75.
- Maiden, Martin (2008). 'Haunting and exorcism: autonomous morphology and phonological conditioning in the Romance (and especially the Romansh) verb.' Paper presented at the *Fourth Oxford-Kobe Linguistics Seminar: History and Structure of the Romance Languages*, Kobe, Japan, April 2008.
- Maiden, Martin (2009). 'Where does heteroclisis come from? Evidence from Romanian dialects.' *Morphology* 19(1): 59-86.
- Maiden, Martin (2011a). 'Allomorphy, autonomous morphology and phonological conditioning in the history of the Daco-Romance present and subjunctive.' *Transactions of the Philological Society* 109(1): 59-91.
- Maiden, Martin (2011b). 'Morphemes and "stress-conditioned allomorphy" in Romansh', in M. Maiden et al. (eds.), 36-50.
- Maiden, Martin, Smith, John Charles, Goldbach, Maria, and Hinzelin, Marc-Olivier (eds.) (2011). *Morphological autonomy: perspectives from Romance inflectional morphology*. Oxford: Oxford University Press.
- Marin, Maria and Marinescu, Bogdan (1984). 'Graiurile din Transilvania', in V. Rusu (ed.), *Tratat de dialectologie românească*. Craiova: Scrisul Românesc, 354-90.
- Martzloff, Vincent (2006). *Les thèmes de présent en yod dans l'épigraphie italique et en latin archaïque*. Doctoral thesis, Université Lumière – Lyon II.
- Matthews, Peter H. (1972). *Inflectional morphology: a theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.

- Matthews, Peter H. (2007). *The concise Oxford dictionary of linguistics*. 2nd edition. Oxford: Oxford University Press.
- Mavrochalyvidis, Georgios and Kesisoglou, Ioannis I. (1960). *Τὸ γλωσσικὸ ἰδίωμα τῆς Ἄξου* [Tò glossikò idíoma tís Axoû] (*Le dialecte d'Axos*). Athens: Institut Français d'Athènes.
- Mayerthaler, Willi (1981). *Morphologische Natürlichkeit*. Wiesbaden: Athenäum.
- Mayerthaler, Willi (1987). 'System-independent morphological naturalness', in W. U. Dressler, W. Mayerthaler, O. Panagl, and W. U. Wurzel (eds.), *Leitmotifs in Natural Morphology*. Amsterdam: John Benjamins, 25-58.
- Mâzuc, Émile (1899). *Grammaire languedocienne: dialecte de Pézenas*. Toulouse: Éditions Privat.
- Meier-Brügger, Michael, Fritz, Matthias, and Mayrhofer, Manfred (eds.) (2010). *Indogermanische Sprachwissenschaft*. 9th edition. Berlin: De Gruyter.
- Meillet, Antoine (1908). *Les dialectes indo-européens*. Paris: Champion.
- Meillet, Antoine (1922). 'Latin *procitum*.' *Bulletin de la Société de linguistique de Paris* 23: 81-83.
- Meiser, Gerhard (1986). *Lautgeschichte der umbrischen Sprache*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Meiser, Gerhard (1998). *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Meiser, Gerhard (2003). *Veni vidi vici: die Vorgeschichte des lateinischen Perfektsystems*. Munich: C.H. Beck.
- Menges, Karl H. (1978). 'Problems of Tungus linguistics.' *Anthropos* 73(3/4): 367-400.
- Meul, Claire (2013). *The Romance reflexes of the Latin infixes -i/esc- and -idi-: restructuring and remodeling processes*. Hamburg: Buske.
- Morariu, Leca (1924). 'Morfologia verbului predicativ român.' *Codrul Cosminului* 1: 1-96.
- Mourin, Louis and de Poerck, Guy (1964). *Introduction à la morphologie comparée des langues romanes, basée sur des traductions anciennes des Actes des Apôtres ch. 20 à 24*, vol. 4: *Sursilvain et engadinois anciens, et ladin dolomitique*. Bruges: De Tempel.
- Müller, Gereon (2007). 'Notes on paradigm economy.' *Morphology* 17(1): 1-38.

- Müller, Gereon, Gunkel, Lutz, and Zifonun, Gisela (2004). 'Introduction', in G. Müller, L. Gunkel, and G. Zifonun (eds.), *Explorations in nominal inflection*. Berlin: Mouton de Gruyter, 1-20.
- Nagy, Gregory (1970). *Greek dialects and the transformation of an Indo-European process*. Cambridge, Mass.: Harvard University Press.
- NALRBanat = Beltechi et al. (2005).
- NALRCrișana = I. Stan and D. Urișescu (eds.), *Noul atlas lingvistic român pe regiuni. Crișana*. Unpublished material.
- NALROltenia = Teaha et al. (1984).
- Neagoe, Victorela (1984). 'Subdialectul bănațean', in V. Rusu (ed.), *Tratat de dialectologie românească*. Craiova: Scrisul Românesc, 240-84.
- Nedelcu, Isabela (2013). 'Inflectional classes of verbs', in G. Pană Dindelegan (ed.), *The grammar of Romanian*. Oxford: Oxford University Press, 18-23.
- Neue, Friedrich and Wagener, Carl (1902). *Formenlehre der lateinischen Sprache*, vol. 1: *Das Substantivum*. 3rd edition. Heidelberg: C. Winter.
- Niedermann, Max (1908). 'Une loi rythmique proethnique en latin', in *Mélanges de linguistique offerts à M. Ferdinand de Saussure*. Paris: Société de linguistique de Paris, 43-57.
- Nielsen Whitehead, Benedicte (2012). *Pickpocket compounds from Latin to Romance*. Doctoral thesis, Universiteit Leiden.
- Nikolaeva, Irina and Spencer, Andrew (2008). 'Nouns as adjectives and adjectives as nouns.' Ms., SOAS and University of Essex. Online: privatewww.essex.ac.uk/~spena/PossMod/8adjnoun-language-revised-final-non-anon.pdf
- Nishimura, Kanehiro (2008). *Vowel reduction and deletion in Italic: effects of stress*. Doctoral dissertation, UCLA.
- Noyer, Rolf (2005). 'A constraint on interclass syncretism', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Springer, 273-315.
- Nübling, Damaris (2001). 'Wechselflexion Luxemburgisch-Deutsch kontrastiv: *ech soen - du sees / si seet* vs. *ich sage, du sagst, sie sagt*. Zum sekundären Ausbau eines präsentischen Wurzelvokalwechsels im Luxemburgischen.' *Sprachwissenschaft* 26(4), 433-71.
- Nyman, Martti (1987). 'Is the Paradigm Economy Principle relevant?' *Journal of Linguistics* 23(2): 251-67.
- Nyman, Martti (1988). 'Paradigm economy: a rejoinder to Carstairs.' *Journal of Linguistics* 24(2): 501-13.

- Orza, Rodica (1979). 'Cu privire la iotacizarea verbelor în graiurile limbii române.' *Cercetării de lingvistică* 24: 73-81.
- Palancar, Enrique (2012). 'The conjugational classes of Tilapa Otomi: an approach from canonical typology.' *Linguistics* 50(4): 783-832.
- Papadopoulos, Anthimos A. (1926). *Γραμματική τῶν Βορείων Ἰδιωμάτων τῆς Νέας Ἑλληνικῆς Γλώσσης* [Grammatikì tòn Voreíon Idiómáton tîs Néas Ellinikîs Glóssis]. Athens: P. D. Sakellarios.
- Penny, Ralph (2002). *A history of the Spanish language*. Cambridge: Cambridge University Press.
- Petersson, Herbert (1921). *Studien über die indogermanische Heteroklisie*. Lund: Gleerup.
- Pinkster, Harm (1990). *Latin syntax and semantics*. London: Routledge.
- Plank, Frans (1991). 'Of abundance and scantiness in inflection: a typological prelude', in F. Plank (ed.), *Paradigms: the economy of inflection*. Berlin/New York: Mouton de Gruyter, 1-39.
- Plank, Frans (1996). 'Patterns of suppletion in inflection.' Paper presented at the *Seventh International Morphology Meeting*, University of Vienna, 16th-18th February 1996.
- Pop, Sever (1952). 'La iotacisation dans les langues roumaines', in *Mélanges de linguistique et de littérature romanes offerts à Mario Roques*, vol. 3. Paris: Bado, 195-235.
- Pope, Mildred K. (1956). *From Latin to modern French with especial consideration of Anglo-Norman: phonology and morphology*. Manchester: Manchester University Press.
- Prader-Schucany, Silvia (1970). *Romanisch Bünden als selbständige Sprachlandschaft*. Bern: Francke.
- Probert, Philomen (2003). *A new short guide to the accentuation of Ancient Greek*. London: Bristol Classical Press.
- Ralli, Angela (2006). 'On the role of allomorphy in inflectional morphology: evidence from dialectal variation', in G. Sica (ed.), *Open problems in linguistics and lexicography*. Monza: Polimetrica, 123-51.
- Ravier, Xavier (1978-93). *Atlas linguistique et ethnographique du Languedoc occidental* [4 vols]. Paris: Éditions du CNRS.
- REW = Meyer-Lübke, Wilhelm (1935). *Romanisches etymologisches Wörterbuch*. 3rd edition. Heidelberg: C. Winter.

- Ringe, Donald (2006). *A linguistic history of English*, vol. 1: *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.
- Risch, Ernst (1977). 'Das System der lateinischen Deklinationen.' *Cahiers Ferdinand de Saussure* 31: 229-45.
- Rix, Helmut (1965). 'Lat. *iecur*, *iocineris*.' *Münchener Studien für Sprachwissenschaft* 18: 79-92.
- Rix, Helmut (1992). 'Zur Entstehung des lateinischen Perfektparadigmas', in O. Panagl and T. Krisch (eds.), *Latein und Indogermanisch. Akten des Kolloquiums der Indogermanischen Gesellschaft, Salzburg, 23.-26. September 1986*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft, 221-40.
- Rix, Helmut (1994). 'Latein und Sabellisch: Stammbaum und/oder Sprachbund?' *Incontri Linguistici* 17: 13-29.
- Rix, Helmut (1999). 'Schwach charakterisierte lateinische Präsensstämme zu Setz-Wurzeln mit Vollstufe I', in H. Eichner and H. C. Luschützky (eds.), *Compositiones indogermanicae in memoriam Jochem Schindler*. Prague: Enigma Corporation, 515-35.
- Rix, Helmut (ed.) (2001). *Lexikon der indogermanischen Verben*. 2nd edition. Wiesbaden: Reichert.
- Romieu, Maurice and Bianchi, André (2005). *Gramatica de l'occitan gascon contemporanèu*. Bordeaux: Presses Universitaires de Bordeaux.
- Ronjat, Jules (1937). *Grammaire istorique des parlers provençaux modernes*, vol. 3: *Morphologie et formation des mots. Notes et syntaxe*. Montpellier: Société des Langues Romanes.
- Rosetti, Alexandru (1968). *Istoria limbii române de la origini pînă în secolul al XVII-lea*. 3rd edition. Bucharest: Editura pentru Literatură.
- Round, Erich R. (2012). *Kayardild morphology and syntax*. Oxford: Oxford University Press.
- Rubenbauer, Hans and Hofmann, Johann B. (1995). *Lateinische Grammatik* [12th edition, revised by Rolf Heine]. Bamberg: Buchner.
- Sagot, Benoît and Walther, Géraldine (2011). 'Non-canonical inflection: data, formalisation and complexity measures', in C. Mahlow and M. Piotrowski (eds.), *Systems and frameworks for computational morphology: second international workshop, SFCM 2011, Zürich, Switzerland, August 26, 2011. Proceedings*. Berlin: Springer, 23-45.
- Sala, Marius (1964). 'r și rr în secolul al XVI-lea.' *Studii și cercetări lingvistice* 15: 170-79.

- Sampson, Rodney (1999). *Nasal vowel evolution in Romance*. Oxford: Oxford University Press.
- Sampson, Rodney (2010). *Vowel prosthesis in Romance: a diachronic study*. Oxford: Oxford University Press.
- Sánchez Miret, Fernando (2006). 'Productivity of the weak verbs in Romanian.' *Folia Linguistica* 40(1/2): 29-50.
- Saramandu, Nicolae (2005). *Atlasul lingvistic român pe regiuni: sinteză*, vol. 1. Bucharest: Editura Academiei Române.
- Sauzet, Patrick and Ubaud, Josiane (1995). *Lo vèrb occitan : guide complet de conjugaison selon les parlers languedociens*. Aix-en-Provence: Édisud.
- Schad, Samantha (2007). *A lexicon of Latin grammatical terminology*. Pisa/Rome: Fabrizio Serra Editore.
- Schindler, Jochem (1975). 'L'apophonie des thèmes indo-européens en -r/n.' *Bulletin de la Société de linguistique de Paris* 70: 1-10.
- Schmid, Heinrich (1982). *Richtlinien für die Gestaltung einer gesamtbündnerromanischen Schriftsprache: Rumantsch Grischun*. Chur: Lia Rumantscha.
- Schorta, Andrea (1938). *Lautlehre der Mundart von Müstair: mit Ausblicken auf die sprachlichen Verhältnisse des inneren Münstertals*. Paris: Droz.
- Schrijver, Peter (1991). *The reflexes of the Proto-Indo-European laryngeals in Latin*. Amsterdam: Rodopi.
- Schrijver, Peter (2003). 'Athematic *i*-Presents: the Italic and Celtic evidence.' *Incontri Linguistici* 26: 59-86.
- Schulte, Kim (2005). 'Vowel centralization in Romanian verbs of Slavic origin: deliberate exploitation of an indigenous sound change?', in T. Geerts, I. van Ginneken, and H. Jacobs (eds.), *Romance languages and linguistic theory 2003*. Nijmegen: Radboud University, 311-25.
- Schumacher, Stefan (2004). *Die keltischen Primärverben*. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck.
- Seldeslachts, Herman (1997). *Étude historico-comparative des racines ultimae laryngalis dans le système verbal latin. Contribution à la morphologie indo-européenne*. Doctoral thesis, Katholieke Universiteit Leuven.
- Seldeslachts, Herman (2001). *Études de morphologie historique du verbe latin et indo-européen*. Leuven: Peeters.

- Sen, Ranjan (2012). 'Exon's Law and the Latin syncope', in P. Probert and A. Willi (eds.), *Laws and rules in Indo-European*. Oxford: Oxford University Press, 205-26.
- Shannon, Claude (1948). 'A mathematical theory of communication.' *The Bell System Technical Journal* 27: 379-423, 623-56.
- Shevelov, George Y. (1964). *A prehistory of Slavic: the historical phonology of common Slavic*. Heidelberg: C. Winter.
- Signorell, Faust, Wüthrich-Grisch, Mena, and Simeon, Gion Pol (1987). *Normas surmiranas. Grammatica rumantscha digl idiom da Sur- e Sotses*. Chur: Tgesa editoura cantunala per meds d'instrucziun.
- Sihler, Andrew L. (1995). *New comparative grammar of Greek and Latin*. Oxford: Oxford University Press.
- Simon, Horst J. and Wiese, Heike (eds.) (2011). *Expecting the unexpected: exceptions in grammar*. Berlin: Mouton de Gruyter.
- Sims, Andrea D. (2007). 'Why defective paradigms are, and aren't, the result of competing morphological patterns', in M. Elliott, J. Kirby, O. Sawada, E. Staraki, and S. Yoon (eds.), *CLS 43: the panels. Papers from the forty-third annual meeting of the Chicago Linguistic Society*, vol. 2. Chicago: Chicago Linguistic Society, 267-81.
- Skalička, Vladimir (1979). *Typologische Studien*. Braunschweig: Vieweg.
- Sommer, Ferdinand (1914). *Handbuch der lateinischen Laut- und Formenlehre: eine Einführung in das sprachwissenschaftliche Studium des Lateins*. 2nd and 3rd edition. Heidelberg: C. Winter.
- Spencer, Andrew (2007). 'Extending deponency', in M. Baerman et al. (eds.), 45-70.
- Spencer, Andrew (2013). *Lexical relatedness: a paradigm-based model*. Oxford: Oxford University Press.
- Spescha, Arnold (1989). *Grammatica sursilvana*. Chur: Casa editura per mieds d'instrucziun.
- Stampe, David (1969). 'The acquisition of phonetic representation', in R. I. Binnick, A. Davidson, G. M. Green, and J. L. Morgan (eds.), *Papers from the fifth regional meeting of the Chicago Linguistic Society*. Chicago: Department of Linguistics, University of Chicago, 443-54.
- Stewart, Thomas and Stump, Gregory T. (2007). 'Paradigm function morphology and the morphology-syntax interface', in G. Ramchand and C. Reiss (eds.), *The Oxford handbook of linguistic interfaces*. Oxford: Blackwell, 383-421.

- Stimm, Helmut and Linder, Karl P. (1989). 'Bündnerromanisch: Interne Sprachgeschichte I. Grammatik', in G. Holtus, M. Metzeltin, and C. Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. 3: *Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart. Rumänisch, Dalmatisch/Istroromanisch, Friaulisch, Ladinisch, Bündnerromanisch*. Tübingen: Niemeyer, 764-85.
- Stump, Gregory T. (2001). *Inflectional morphology: a theory of paradigm structure*. Cambridge: Cambridge University Press.
- Stump, Gregory T. (2005a). 'Referrals and morphemes in Sora verb inflection', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2005*. Dordrecht: Springer, 227-51.
- Stump, Gregory T. (2005b). 'Some criticisms of Carstairs-McCarthy's conclusions', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 2005*. Dordrecht: Springer, 283-303.
- Stump, Gregory T. (2006). 'Heterocclisis and paradigm linkage.' *Language* 82(2): 279-322.
- Stump, Gregory T. and Finkel, Raphael A. (2013). *Morphological typology: from word to paradigm*. Cambridge: Cambridge University Press.
- Szemerényi, Oswald J. L. (1960). 'Etyma Latina I (1-6).' *Glotta* 38(3/4): 216-51.
- Szemerényi, Oswald J. L. (1996). *Introduction to Indo-European linguistics*. Oxford: Clarendon.
- Tanaka, Toshiya (2011). *A morphological conflation approach to the historical development of preterite-present verbs: Old English, Proto-Germanic, and Proto-Indo-European*. Fukuoka: Hana-Shoin.
- Taylor, Daniel J. (1991). 'Latin declensions and conjugations: from Varro to Priscian.' *Histoire, Epistémologie, Langage* 13(2): 85-109.
- Teaha, Teofil, Ionică, Ion, and Rusu, Valeriu (eds.) (1984). *Noul atlas lingvistic român pe regiuni. Oltenia*, vol. 5. Bucharest: Editura Academiei Române.
- Tekavčić, Pavao (1972/3 & 1974). 'Abbozzo del sistema morfosintattico del soprasilvano odierno.' *Studia Romanica et Anglica Zagrabiensia* 33-6: 359-488; 37: 5-134.
- Thomas, François (1935). 'Sur les désidératifs latins en -essō.' *Revue de philologie (3rd series)* 9(3): 280-87.
- Thöni, Gion P. (1964). *Rumantsch-Surmeir: grammatica per igl idiom surmiran*. Chur: Ligia Romontscha.

- Thornton, Anna Maria (2011). 'Overabundance (multiple forms realizing the same cell): a non-canonical phenomenon in Italian verb morphology', in M. Maiden et al. (eds.), 358-81.
- Thurneysen, Rudolf (1879). *Über Herkunft und Bildung der lateinischen Verba auf -io der dritten und vierten Conjugation und über ihr gegenseitiges Verhältniss*. Leipzig: J.B. Hirschfeld.
- Tiktin, Hariton (1905). *Rumänisches Elementarbuch*. Heidelberg: C. Winter.
- Touratier, Christian (1994). *Syntaxe latine*. Leuven: Peeters.
- Tscharner, Gion (2012). *Verbs valladers*. 4th edition. Chur: Südostschweiz Print.
- Untermann, Jürgen (2000). *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: C. Winter.
- Uricaru, Lucia and Goga, Mircea (1997). *Verbes roumains*. Cluj: Editura Echinox.
- Vieli, Ramun and Decurtins, Alexi (1962). *Vocabulari romontsch sursilvan-tudestg*. Chur: Ligia Romontscha.
- von Planta, Robert (1892). *Grammatik der oskisch-umbrischen Dialekte*, vol. 1: *Einleitung und Lautlehre*. Strassburg: Trübner.
- von Planta, Robert and Melcher, Florian (eds.) (1939-). *Dicziunari rumantsch grischun*. Winterthur: Società Retorumantscha.
- Wade, Terence (1992). *A comprehensive Russian grammar*. Oxford: Blackwell.
- Walther, Géraldine (2011). 'Latin passive morphology revisited.' Paper presented at the *Annual Meeting of the Linguistics Association of Great Britain*, University of Manchester, 7th-10th September 2011.
- Watkins, Calvert (1969). *Indogermanische Grammatik*, vol. 3/1: *Geschichte der indogermanischen Verbalflexion*. Heidelberg: C. Winter.
- Watkins, Calvert (1971). 'Hittite and Indo-European studies: the denominative statives in *-ē-.' *Transactions of the Philological Society* 70: 51-93.
- Weiss, Michael L. (1996). 'Review of Sihler (1995).' *American Journal of Philology* 117(4): 670-75.
- Weiss, Michael L. (2009). *Outline of the historical and comparative grammar of Latin*. Ann Arbor: Beech Stave Press.
- Winzap, Isidor (1986). *La conjugaziun dils verbs romantschs sursilvans*. Chur: Casa editura cantunala per mieds d'instrucziun.

- Winzap, Isidor (1990). 'Aspects e problems dalla conjugaziun dils verbs romontschs sursilvans.' *Annalas da la Societad Retorumantscha* 103: 29-41.
- Wurzel, Wolfgang U. (1984). *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung*. Berlin: Akademie-Verlag.
- Wurzel, Wolfgang U. (1989). 'Review of: Andrew Carstairs, *Allomorphy in Inflection*', in G. E. Booij and J. van Marle (eds.), *Yearbook of Morphology 1989*. Dordrecht: Foris, 205-27.
- Wurzel, Wolfgang U. (1990). 'The mechanism of inflection: lexicon representations, rules, and irregularities', in W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer, and J. R. Rennison (eds.), *Contemporary morphology*. Berlin/New York: Mouton de Gruyter, 203-16.
- Zafiu, Rodica (2013). 'Mood, tense and aspect', in G. Pană Dindelegan (ed.), *The grammar of Romanian*. Oxford: Oxford University Press, 24-65.
- Zamfir, Dana-Mihaela (2005). *Morfologia verbului în dacoromâna veche (secolele al XVI-lea – al XVII-lea)*, vol. 1. Bucharest: Editura Academiei Române.
- Zapheiriou, Menekratis P. (1995). *To γλωσσικό ιδίωμα της Σάμου* [To glossikó idíoma tis Sámou]. Athens: Grigoris.
- Zwicky, Arnold M. (1985). 'How to describe inflection', in M. Niepokuj, M. Van Clay, V. Nikiforidou, and D. Feder (eds.), *Proceedings of the eleventh annual meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 372-86.